EasyChair Preprint

№ 11102

# HetBiSyn: Predicting Anticancer Synergistic Drug Combinations Featuring Bi-Perspective Drug Embedding with Heterogeneous Data

Yulong Li, Hongming Zhu, Xiaowen Wang and Qin Liu



October 23, 2023

# HetBiSyn: Predicting Anticancer Synergistic Drug Combinations Featuring Bi-perspective Drug Embedding with Heterogeneous Data

Yulong Li[1[0000−0002−3068−2554]], Hongming Zhu[1[0000−0001−5795−5279]], Xiaowen Wang[1[0000−0003−1880−7921]], and Qin Liu[1[0000−0002−9352−1694]⋆]

[1]School of Software Engineering, Tongji University, Shanghai, China

**Abstract.** Synergistic drug combination is a promising solution to cancer treatment. Since the combinatorial space of drug combinations is too vast to be traversed through experiments, computational methods based on deep learning have shown huge potential in identifying novel synergistic drug combinations. Meanwhile, the feature construction of drugs has been viewed as a crucial task within drug synergy prediction. Recent studies shed light on the use of heterogeneous data, while most studies make independent use of relational data of drug-related biomedical interactions and structural data of drug molecule, thus ignoring the intrinsical association between the two perspectives. In this study, we propose a novel deep learning method termed HetBiSyn for drug combination synergy prediction. HetBiSyn innovatively models the drug-related interactions between biomedical entities and the structure of drug molecules into different heterogeneous graphs, and designs a self-supervised learning framework to obtain a unified drug embedding that simultaneously contains information from both perspectives. In details, two separate heterogeneous graph attention networks are adopted for the two types of graph, whose outputs are utilized to form a contrastive learning task for drug embedding that is enhanced by hard negative mining. We also obtain cell line features by exploiting gene expression profiles. Finally HetBiSyn uses a DNN with batch normalization to predict the synergy score of a combination of two drugs on a specific cell line. The experiment results show that our model outperforms other state-of-art DL and ML methods on the same synergy prediction task. The ablation study also demonstrates that our drug embeddings with bi-perspective information learned through the end-to-end process is significantly informative, which is eventually helpful to predict the synergy scores of drug combinations.

**Keywords:** synergistic drug combinations · Deep learning · Heterogeneous data · Graph attention network · Self-supervised learning

## 1 Introduction

In the field of cancer treatment, drug combination therapy[1] holds significant importance. The interactions of pairwise combinations of drugs can be divided

into synergistic, additive, and antagonistic by comparing the effect of the drug combination with the sum of effects of drugs applied separately[2]. Synergistic drug combinations can often reduce the development of drug resistance[3] and minimize the occurrence of drug-related side effects[4] during the treatment process. However, distinguishing synergistic drug combinations from non-synergistic ones is challenging as the combination space expands rapidly with the discovery of new drugs.

Early studies on synergistic drug combinations are mostly based on clinical experience, which is time-consuming and labor-intensive, and may lead to unnecessary or even harmful treatments on patients[5]. Even applying high-throughput screening technology (HTS)[6] that enables efficient testing of cell lines *in vitro* is impossible to screen through the complete combination space, let alone that the technology is expensive to build[7]. Researchers have therefore turned to computational methods to predict synergistic drug combinations.

Except for computational models that are only available on specific drugs or cell lines[5], recent studies have shed light on methods based on machine learning (ML) and deep learning (DL). The most common workflow consists of obtaining the features for cell lines and drugs and predicting the synergy score with a ML or DL model. Previous studies employed various ML models [8,9,10,11,12] to predict the synergy of anticancer drug combinations. In recent years, the availability of large-scale synergy datasets[13] has provided a valuable resource for employing DL methods in drug combination prediction. Commonly adopted DL models include Deep Neural Network (DNN)[5,2], Residual Neural Network[14], and other interpretable DL models. Additionally, special techniques such as Ensemble Learning[15,16], Transfer Learning[17], and Tensor factorization[18] have been adopted to predicting drug combination synergy.

On the other hand, drug features play an essential role in the synergy prediction task. A classical and universal method for drug representation is to directly use molecular fingerprints[19] or molecular descriptors[5], which refers to a predefined feature vector containing substructural and physicochemical properties. Some researchers[5,2,11,12] collected the interactions between drugs and other biomedical entities (e.g. drug target, pathway etc.), and simply obtained the feature vector by sampling a binary digit for each interaction. Those feature-engineering-based methods offer easy access to fairly informative representations, whereas they might be greatly affected by prior assumptions of biochemical domain knowledge. Methods based on representation learning are proposed to alleviate this problem[17,20,21,22,23].

When it comes to how data is utilized in constructing drug features, it should be emphasized that the chemical structure of a drug determines how it functions, while the interactions between drugs and other biomedical entities represent known patterns of drug action. Both aspects should be considered for comprehensive drug features. Concerning the way that molecular-level data and drug-related interaction are used in drug feature constructing, previous studies on synergy prediction either use only one type of data or simply perform concatenation after representations from both aspects are extracted respectively. As

the two aspects are intrinsically associated but differs a lot in terms of data, such methods might not be able to fully exploit the latent information when the features are directly used for subsequent synergy prediction. There is still potential for improvement by devising a drug representation learning model to simultaneously obtain information from both structural and relational heterogeneous data.

To improve the synergy prediction of anti-cancer drug combination from the perspective of constructing compact drug representations that are more expressive and informative, we hereby propose a deep learning model **HetBiSyn** (Drug **Syn**ergy Prediction featuring **Bi**-perspective Drug Embedding with **Het**erogeneous Data). In this paper, drug embeddings that integrate information from both aspects are learned within a self-supervised training process. Specifically, HetBiSyn constructs a graph for interactions between drug-related biomedical entities and multiple molecular-level graphs for drugs. Both types of graphs are heterogeneous, and separate Heterogeneous Graph Attention Networks (HGAT) are applied to each type of graph to embed information from different perspective. A contrastive learning module is designed to learn a unified embedding based on the output of the two HGATs, and the hard sample mining strategy is adopted to enhance the model. Besides, HetBiSyn utilizes gene expression profiles to construct the cell line features. Lastly, a DNN with Batch Normalization mechanism is designed to predict the synergy score of drug combinations on cell lines. We compared HetBiSyn with other popular ML and DL methods on the synergy dataset contributed by O'Neil, and the result demonstrates that HetBiSyn can achieve more accurate drug synergy prediction.

## 2 Materials and Methods

### 2.1 Synergy Dataset

A high-throughput drug combination screening dataset was obtained from O'Neil's research. The dataset encompass 583 pairwise drug combinations involving 38 distinct drugs tested against 39 human cancer cell lines. Preuer et al.[5] computed a synergy score for each sample using Loewe Additivity values, and divided all samples into 5 disjoint folds with an equal count of drug combinations.

### 2.2 Cell Line Features

The cell line features are extracted mainly based on the gene expression data. The gene expression files are fetched from the ArrayExpress database[24] (accession number: E-MTAB-3610). We adopted the Factor Analysis for Robust Microarray Summarization method[25] to implement quantile normalization and summarization on the gene expression data. The method also provides calls on whether a gene is informative, by which effective genes are selected for the feature construction of cell lines. In all, 3739 genes are screened out and z-score normalization is performed to produce the feature vector.

### 2.3 Construction of Drug-related Graphs

To directly exploit information from different perspectives, we design two types of drug-related graphs from which the embeddings of drugs are learned jointly. The bioinformatic graph provides identified patterns of drug action by integrating the interactions between drugs and other biomedical entities, while molecular-level graphs reveal the structural and chemical particulars inside a drug molecule.

**Drug-related Heterogeneous Bioinformatic Graph** Bioinformatic graphs, also called bioinformatic networks, are widely used in various drug-related problems, especially in extracting complex hidden features that implicate proven patterns of drug actions. Here we construct a heterogeneous graph $G_{Bio} = (V, E)$ , in which each node $v$ in the node set $V$ belongs to a biomedical entity type $t_v$ in a type set $T$, and each edge $e$ in the edge set $E$ belongs to a relation type $t_e \in S_t \times S_t$. As defined, there is at most 1 edge between 2 nodes, and all edges are set bidirectional in practice. Only the largest connected component of $G_{bio}$ is retained so that information can be propagated through every single node. Also, we collect the biomedical entities and their relationships from Luo et al.'s work[26] and DrugBank (Version 3.0), and supplement the drug-target-interaction data with UniProtKB so that all 38 drugs in the synergy dataset are involved in $G_{bio}$.

**Heterogeneous Molecular-level Graph** To exploit drug properties from a microscopic perspective, we construct a graph $G_{mol}$ for **each** drug in $G_{bio}$ at molecular level. First, a molecular graph $G_{mg}$ is generated for each drug by treating the atoms as nodes and the bonds between them as edges. All atom nodes are considered to be of the same type though they are initialized with different atomic features (e.g. chirality, formal charge, partial charge, etc.)[27], while edges vary in types according to the original bond types (e.g. single, double, aromatic, etc.). The edges in $G_{mg}$ are bidirectional since chemical bonds are unbiased. We use the RDKit tool to convert the SMILES string of a drug into a molecule object for subsequent operations, and drugs that do not have a SMILES or cannot be converted are abandoned from $G_{bio}$. Inspired by Fang[28], we augment $G_{mg}$ to leverage the associations between atoms that are not directly connected with bonds but share fundamental chemical attributes. By histogramizing the continuous attributes of atoms and converting them into discrete labels, totally 107 attributes of atoms and 17 relation types are devised. These attributes are then added to $G_{mol}$ as another type of nodes, while their relations with the atoms in $G_{mg}$ are modeled as different types of directed edges pointing to atoms.

### 2.4 HetBiSyn

In this paper, a novel deep learning method named HetBiSyn is proposed to predict synergy scores of drug combinations on cell lines. The overview of our method is shown in Figure 1.

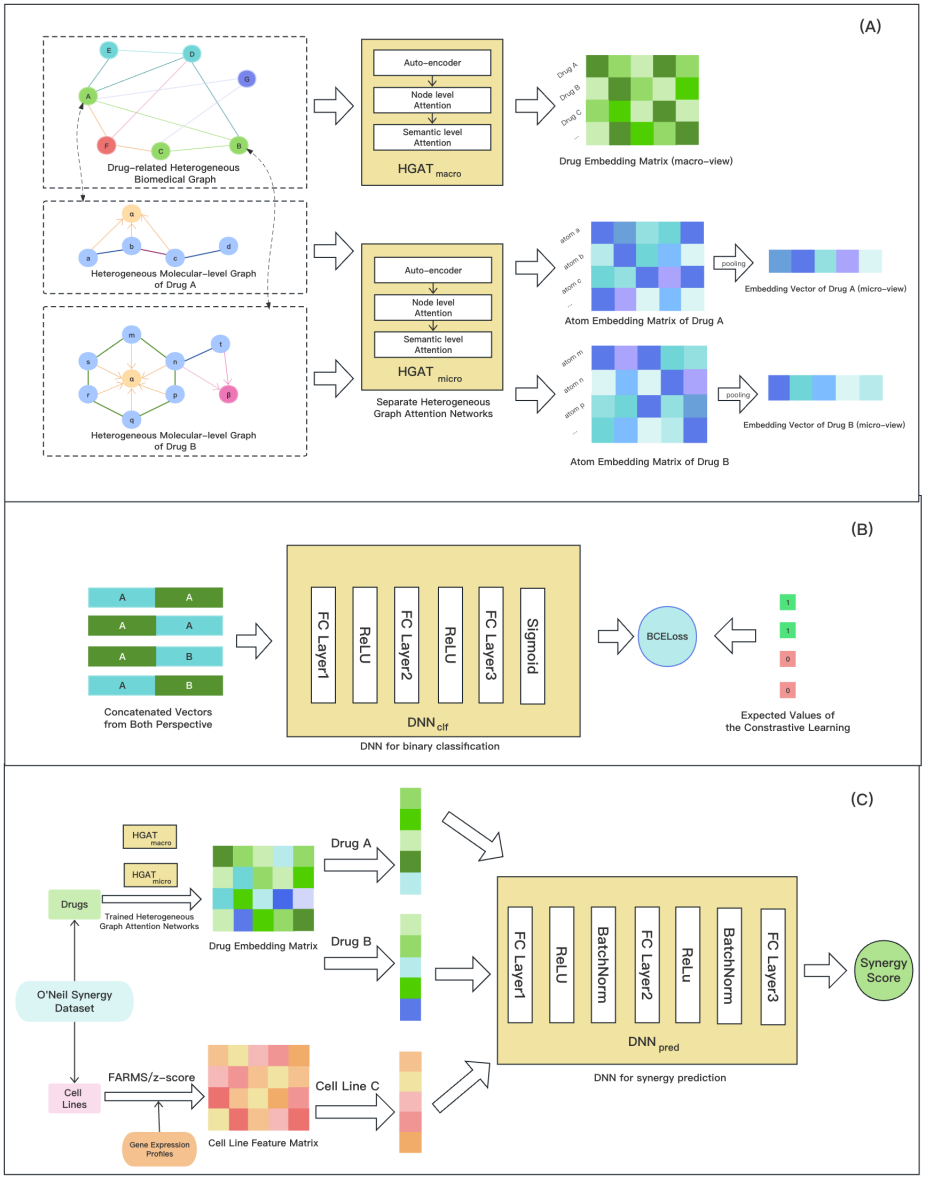


**Fig. 1.** Overview of HetBiSyn. (A) Two HGATs as shown integrate information from the drug-related heterogeneous biomedical graph $G_{bio}$ and the heterogeneous molecular-level graphs $G_{mol}$ respectively, yielding drug embedding at different perspectives. (B) Embedding of drugs from different perspectives are paired to form a contrastive learning task, where $DNN_{clf}$ is set up for binary classification. (C) HetBiSyn predicts the synergy score of a drug combination on a specific cell line using the DNN with the batch normalization mechanism.

**Heterogeneous Graph Attention Networks for Drugs** As shown in figure(B), the essential step for drug representation learning is to extract an intermediate embedding from both $G_{bio}$ and $G_{mol}$ for a drug, which would be continuously updated throughout the subsequent self-supervised learning process. Heterogeneous graph attention network (HGAT)[29] is a node representation learning model that can generate dense embedding while retaining information about network topology and meta-path importance with insights into heterogeneity. HetBiSyn set up two separate HGATs for $G_{bio}$ and $G_{mol}$ to exploit inter-entity and intra-molecular information, namely $HGAT_{macro}$ and $HGAT_{micro}$. The detailed derivation about how HGAT works within our study can be found in supplementary file (section 1).

It is worth noting is that only the embedding for each atom or property node is obtained through training $HGAT_{micro}$, which cannot be directly referred to individual drugs. In order to present the drug embedding by micro-view, average pooling is conducted on the atom node vectors for each drug as whole-graph embedding.

**Drug Representation Learning Based on Contrastive Learning** The subsequent step of drug embedding learning is to leverage the output of the two networks to train a unified embedding balancing both perspective. The key idea comes that the representations of the same drug generated from the two networks shall be as similar as possible, while drugs showing great distinction shall have differentiated embeddings. Under this assumption, we form a binary contrastive learning task aiming at estimating the performance of the outputs and optimizing the recurrent training process.

Let $Z(G, d_i)$ denote the output embedding of drug $d_i$ from either HGAT. For each drug $d_i$ in $G_{bio}$, $S(d_i, d_i) = [Z(HGAT_{macro}, d_i)||Z(HGAT_{micro}, d_i)]$ is defined as a positive sample that is labeled 1, while $S(d_i, d_j)$ given $i \neq j$ is defined as a negative sample that is labeled 0. The concatenated vector $S$ is sampled twice in reverse order with respect to the two networks to generate both positive and negative samples. A simple deep neural network, denoted as $DNN_{clf}$, is set up for binary classification. We use binary cross-entropy loss with sum as reduction for the loss function. The loss function can be described as:

$$loss = \sum_{i=1}^{n} [y_i \cdot log(p_i) + (1 - y_i) \cdot log(1 - p_i)]; p_i = \frac{1}{1 + e^{-x_i}} \tag{1}$$

where $x_i$ and $y_i$ respectively represents the predicted label value and the true label value of a sample. It is worth noting that the drug embedding learning module is an end-to-end process, as the loss computed here is used to update $DNN_{clf}$ as well as the two HGATs through back propagation. After multiple rounds of training, the representation of a drug can be inferred by averaging the output of its embedding from both HGATs.

Furthermore, the sampling strategy is also improved by hard negative mining. In other words, we try to find drugs that are alike and challenge the classifier

to label their combination correctly. We collected a 167-dimension MACCS fingerprint, which is often used to assess the similarity between molecules, for all the drugs in $G_bio$, and calculated the Tanimoto similarity for each pair of drugs. A drug $d_j$ having the most similarity with another drug $d_i$ indicates this duo may compose a hard negative sample, while other randomly taken samples are viewed as common negative samples. We take positive samples, hard negative samples and common negative samples at a certain ratio for training $DNN_{clf}$.

**Synergy Score Prediction** After the cell line features and drug embeddings are obtained, a regression model is designed for predicting cell-line-specific synergy scores of drug combinations. The input vector is constructed by sequentially concatenating the feature or embedding vectors of the two drugs and the cell line in a data point from the synergy dataset, and the synergy score of this trio serves as the output. Each trio is sampled twice in terms of the input vector by exchanging the order of drug vectors, since the network should not differentiate between permutations of two drugs. The prediction model is implemented by a feed forward DNN composed of three fully connected (FC) layers and two batch normalization layers in between, denoting as $DNN_{pred}$. The number of neurons in each FC layer is funnel-shaped as we have the most neurons in the first FC layer and one neuron representing the predicted synergy score in the last FC layer. ReLU is used as the activation function for the first two FC layers. $DNN_{pred}$ takes mean square error loss as its loss function.

## 3 Result

### 3.1 Experiment Setup

We resort to the mean square error (MSE) and the root mean square error (RMSE) for the main evaluation metric. The Pearson correlation coefficient (PCC) between the predictions and the ground truth is also adopted. As the experiments are conducted under a 5-fold cross-validation approach, we present the mean and the standard deviation of each evaluation metric across the 5-fold dataset. Hyper-parameter settings are documented in the supplementary file (section 2).

### 3.2 Performance Comparison with Other Models

To objectively appraise the performance of HetBiSyn, we compare HetBiSyn with some representative models on the same synergy dataset with a 5-fold cross validation. Four ML methods including Elastic Net[30], Support Vector Regression (SVR)[31], Random Forest[32] and XGBoost[33] are adopted by using the same drug and cell line features for input. We implemented these methods with sklearn and retained the default hyper-parameters. We also selected 4 DL methods for comparison, including DeepSynergy[5], MatchMaker[34], AuDNNSynergy[2]

and DFFNDDS[35]. We adopted their feature generating process and validate them upon our dataset, as described in the supplementary file (section 3).

Results of the experiment of the comparison among the methods above are shown in Table 1. The best and second best performance are shown in bold and italic respectively. HetBiSyn achieves the lowest MSE of 225.90 among all compared methods, which is 11.58% less than DeepSynergy, 6.31% less than the AuDnnSynergy which achieves the second lowest and 20.90% less than XGBoost. The PCC of HetBiSyn, which is the second best across all methods compared, also shows a strong correlation between the model's prediction and the ground truth. The result demonstrates the advantage of HetBiSyn on the synergy prediction task, and the possible reasons are: 1)Drug feature is more informative under the end-to-end self-supervised learning framework of HetBiSyn that integrates data from multiple aspects. 2) The DNN for final prediction may identify the nonlinear patterns in the synergy dataset better in comparison with other DL methods.

**Table 1.** Results of method comparison on the synergy score prediction task

| Type | Method | MSE | RMSE | PCC |
|---|---|---|---|---|
| DL | HetBiSyn | **225.90±31.90** | **15.00±1.02** | *0.74±0.03* |
| | DeepSynergy | 255.49 | 15.91±1.56 | 0.73±0.04 |
| | MatchMaker | 254.37±37.70 | 15.93±1.17 | 0.68±0.03 |
| | AuDNNSynergy | *241.12±43.52* | *15.46±1.44* | **0.74±0.04** |
| | DFFNDDS | 242.37±34.21 | 15.53±1.07 | 0.76±0.02 |
| ML | Elastic Net | 407.06±48.23 | 20.18±1.33 | 0.47±0.03 |
| | SVR | 338.57±53.39 | 18.40±1.48 | 0.58±0.03 |
| | Random Forest | 312.75±44.01 | 17.68±1.13 | 0.61±0.02 |
| | XGBoost | 285.60±44.40 | 17.16±1.31 | 0.68±0.02 |

### 3.3 Ablation Study

To further inspect how the use of heterogeneous data from different perspective contributes to the prediction result in our model, a series of variants of HetBiSyn are designed for comparison mainly by altering the drug feature construction process as:

•**HetBiSyn-Bio** Only the data of drug-related biomedical interactions are utilized. Metapath2vec[36] is applied on the $G_{bio}$ we constructed to extract drug features. We use the implementation provided by DGL [37].

•**HetBiSyn-MolHGT** Only the data of molecular structure of drugs are utilized. We adopt a molecular representation learning framework MolHGT[38] to obtain features from heterogeneous graphic data, which treat atoms and bonds as different types of nodes and edges to extract representations.

•**HetBiSyn-MG** Similar to the original HetBiSyn but the molecular graph $G_{mg}$ mentioned in 2.3 is used instead. Atom properties are not considered so that the structural information completely emerges from the molecular graph.

•**HetBiSyn-Concat** Drug features are constructed by concatenating the representation from HetBiSyn-Bio and HetBiSyn-MG.

All the variants are set to generate a drug feature of 128 dimensions except for HetBiSyn-Concat which is doubled by concatenating. The DNN for synergy prediction is same for each variant. The results are displayed in Table 2 .

**Table 2.** Results of the ablation study

| Method | MSE | RMSE | PCC |
|---|---|---|---|
| HetBiSyn | **225.90±31.90** | **15.00±1.02** | **0.74±0.03** |
| HetBiSyn-Bio | 246.75±48.02 | 15.71±1.52 | 0.68±0.02 |
| HetBiSyn-MolHGT | 248.24±41.91 | 15.76±1.42 | 0.67±0.02 |
| HetBiSyn-MG | 229.98±35.22 | 15.17±1.08 | 0.73±0.03 |
| HetBiSyn-Concat | 236.93±40.05 | 15.39±1.35 | 0.70±0.02 |

HetBiSyn-MG is designed to be a fair comparison to HetBiSyn-MolHGT, as they both depend on heterogeneous molecular graphs in terms of obtaining structural information. HetBiSyn-MG performs better because information from the macro-view perspective is also considered. It can be inferred that using data from either single perspective would not make better performance than integrating them even using simple concatenation. Furthermore, methods of fusing data from the micro-view and macro-view also affect the result of prediction. Though HetBiSyn-Concat outperforms other variants based on single perspective data, the original HetBiSyn shows an advantage to it even having less feature dimensions, which proves that the end-to-end self-supervised learning process of drug feature may better integrate the hidden information from both perspective.

## 4 Conclusion

In this paper, we propose a new DL based method for predicting anti-cancer synergistic drug combinations named HetBiSyn. HetBiSyn models the drug-related interactions between biomedical entities and the structure of drug molecules into different heterogeneous graphs, and designs a self-supervised learning framework to obtain a unified drug embedding that simultaneously contains information from both perspective. In details, two separate heterogeneous graph attention networks are adopted for the two types of graph, whose outputs are utilized to form a contrastive learning task for drug embedding that is enhanced by hard negative mining. We also obtain cell line features by exploiting gene expression profiles. Finally HetBiSyn uses a DNN with batch normalization to predict the synergy score of a combination of two drugs on a specific cell line. The experiment results show that our model outperforms other state-of-art DL and ML methods on the same synergy prediction task. Besides, the ablation study demonstrates that our drug embeddings with bi-perspective information learned through the end-to-end process is significantly informative and expressive, which is helpful to predict the synergy scores of drug combinations.

## References


1. Mokhtari RB, Homayouni TS, Baluch N, et al. Combination therapy in combating cancer. Oncotarget 2017;8(23):38022.
2. Zhang T, Zhang L, Payne PRO, Li F. Synergistic Drug Combination Prediction by Integrating Multiomics Data in Deep Learning Models. Methods Mol Biol. 2021;2194:223-238. doi: 10.1007/978-1-0716-0849-4_12. PMID: 32926369.
3. Liu J, Gefen O, Ronin I, et al. Effect of tolerance on the evolution of antibiotic resistance under drug combinations. Science2020;367(6474):200–4.
4. Kruijtzer C, Beijnen J, Rosing H, et al. Increased oral bioavailability of topotecan in combination with the breast cancer resistance protein and p-glycoprotein inhibitor gf120918. J Clin Oncol 20(13):2943–2950
5. Preuer K, Lewis RP, Hochreiter S, Bender A, Bulusu KC, Klambauer G (2018) Deepsynergy: predicting anti-cancer drug synergy with deep learning. Bioinformatics 34(9):1538–1546
6. Lehár J, Krueger AS, Avery W, et al. Synergistic drug combinations tend to improve therapeutically relevant selectivity. Nat Biotechnol 27(7):659–666
7. Macarron R, Banks MN, Bojanic D et al. Impact of high-throughputscreening in biomedical research. Nat Rev Drug Discov 2011;10:188–95
8. Li X, Yingjie X, Cui H, et al. Prediction of synergistic anti-cancer drug combinations based on drug target network and drug induced gene expression profiles. Artif Intell Med 2017;83:35–43.
9. Li H, Li T, Quang D, et al. Network propagation predicts drug synergy in cancers. Cancer Res 2018;78(18):5446–57.
10. Low YS, Daugherty AC, Schroeder EA, et al. Synergistic drug combinations from electronic health records and gene expression. J Am Med Inform Assoc 2017;24(3):565–76.
11. Jeon M, Kim S, Park S, et al. In silico drug combination discovery for personalized cancer therapy. BMC Syst Biol2018;12(2):59–67.
12. Celebi R, Oliver Bear Don't Walk, Movva R, et al. In-silico prediction of synergistic anti-cancer drug combinations using multi-omics data. Sci Rep 2019;9(1):1–10.
13. O'Neil J, Benita Y, Feldman I, et al. An unbiased oncology compound screen to identify novel combination strategies. Mol Cancer Ther 2016;15(6):1155–62.
14. Xia F, Shukla M, Brettin T, et al. Predicting tumor cell line response to drug pairs with deep learning. BMC Bioinformatics 2018;19(18):71–9.
15. Ding P, Yin R, Luo J, et al. Ensemble prediction of synergistic drug combinations incorporating biological, chemical, pharmacological, and network knowledge. IEEE J Biomed Health Inform 2019;23(3):1336–45.
16. Singh H, Rana PS, Singh U. Prediction of drug synergy score using ensemble based differential evolution. IET Syst Biol 2019;13(1):24–9.
17. Kim Y, Zheng S, Tang J, et al. Anticancer drug synergy prediction in understudied tissues using transfer learning. J Am Med Inform Assoc 2021;28(1):42–51.
18. Sun Z, Huang S, Jiang P, et al. Dtf: deep tensor factorization for predicting anti-cancer drug synergy. Bioinformatics 2020;36(16):4483–9.
19. Rogers D, Hahn M. Extended-connectivity fingerprints. J. Chem. Inf. Model., 50, 742–754.
20. Ekşioğlu I, Tan M. Prediction of Drug Synergy by Ensemble Learning[J]. arXiv:2001.01997, 2020.
21. Bai Y, Gu K, Sun Y, Wang W. Bi-Level Graph Neural Networks for Drug-Drug Interaction Prediction[J]. arXiv:2006.14002, 2020.

22. Winter R, Montanari F, Noé F, Clevert D. Learning Continuous and Data-Driven Molecular Descriptors by Translating Equivalent Chemical Representations[J]. Chemical Science. 10. 10.1039/C8SC04175J, 2019.
23. Yang J, Xu Z, Wu W, et al. GraphSynergy: a network-inspired deep learning model for anticancer drug combination prediction, Journal of the American Medical Informatics Association, Volume 28, Issue 11, November 2021, Pages 2336–2345
24. Iorio F, Knijnenburg TA, Vis DJ, et al. A landscape of pharmacogenomic interactions in cancer. Cell 2016;166(3):740–54.
25. Hochreiter S, Clevert D-A, Obermayer K. A new summarization method for affymetrix probe level data. Bioinformatics 2006;22(8):943–9.
26. Wan F, Hong L, Xiao A, et al. NeoDTI: neural integration of neighbor information from a heterogeneous network for discovering new drug–target interactions, Bioinformatics, Volume 35, Issue 1, January 2019, Pages 104–111
27. Chuang, Kangway V., and Michael J. Keiser. "Comment on 'Predicting Reaction Performance in C–N Cross-coupling Using Machine Learning.'" Science, vol. 362, no. 6416, American Association for the Advancement of Science (AAAS)
28. Fang, Yin, et al. "Molecular Contrastive Learning With Chemical Element Knowledge Graph | Proceedings of the AAAI Conference on Artificial Intelligence." Molecular Contrastive Learning With Chemical Element Knowledge Graph | Proceedings of the AAAI Conference on Artificial Intelligence
29. Wang X, Ji H, Shi C et al. Heterogeneous graph attention network. In:The World Wide Web Conference, WWW '19, pp. 2022–32, NewYork: Association for Computing Machinery, 2019.
30. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Series B Stat Methodology 2005;67(2):301–20.
31. Drucker H, CJC B, Kaufman L, et al. Support vector regression machines. Adv Neural Information Process Syst 1997;9:155–61.
32. Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
33. Chen T, Guestrin C. Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, United States: Association for Computing Machinery (ACM), 2016, 785–94.
34. Halil K, Oznur T, Ercument C. MatchMaker: A Deep Learning Framework for Drug Synergy Prediction. IEEE/ACM Transactions on Computational Biology and Bioinformatics. PP. 1-1. 10.1109/TCBB.2021.3086702.
35. Xu M, Zhao X, Wang J, et al. DFFNDDS: prediction of synergistic drug combinations with dual feature fusion networks. Journal of Cheminformatics. 15. 10.1186/s13321-023-00690-3.
36. Dong Y, Chawla, N. V, Swami A. metapath2vec. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
37. Wang M, Zheng D, Ye Z, et al. Deep Graph Library: A Graph-Centric, Highly-Performant Package for Graph Neural Networks: Vol. arXiv preprint arXiv:1909.01315.
38. Deng D, Lei Z, Hong X, et al. Describe Molecules by a Heterogeneous Graph Neural Network with Transformer-like Attention for Supervised Property Predictions. ACS Omega, 7(4), 3713–3721. https://doi.org/10.1021/acsomega.1c06389