From Cell-Lines to Cancer Patients: Personalized Drug Synergy Prediction

Halil Ibrahim Kuru[1], A. Ercument Cicek[1,2,*] and Oznur Tastan[3,*]

1. Department of Computer Engineering, Bilkent University, Ankara, Turkey 06800
2. Computational Biology Department, Carnegie Mellon University, Pittsburgh, PA 15213
3. Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey 34956
*Correspondence: cicek@cs.bilkent.edu.tr and otastan@sabanciuniv.edu

Abstract. Combination drug therapies are effective treatments for cancer. However, the genetic heterogeneity of the patients and exponentially large space of drug pairings pose significant challenges for finding the right combination for a specific patient. Current *in silico* prediction methods can be instrumental in reducing the vast number of candidate drug combinations. However, existing powerful methods are trained with cancer cell line gene expression data, which limits their applicability in clinical settings. While synergy measurements on cell lines models are available at large scale, patient-derived samples are too few to train a complex model. On the other hand, patient-specific single-drug response data are relatively more available. In this work, we propose a deep learning framework, Personalized Deep Synergy Predictor (PDSP), that enables us to use the patient-specific single drug response data for customizing patient drug synergy predictions. PDSP is first trained to learn synergy scores of drug pairs and their single drug presonses for a given cell line using drug structures and large scale cell line gene expression data. Then, the model is fine-tuned for patients with their patient gene expression data and associated single drug response measured on the patient *ex vivo* samples. In this study, we evaluate PDSP on data from three leukemia patients and observe that it improves the prediction accuracy by 27% compared to models trained on cancer cell line data. PDSP is built and available at https://github.com/hikuru/PDSP

Keywords: Cancer Cell Lines · Multi-task Transfer Learning · Deep Learning · Drug Synergy · Personalized medicine

1 Introduction

Combination therapies, in which multiple drugs are administered concurrently, have emerged as alternatives to single drug therapies for various complex diseases including cancer [1, 18, 20, 23, 26], diabetes [15], human immunodeficiency virus [8], and SARS-CoV-2 [32]. These therapies are reported to provide more effective treatment [9, 29], overcome drug resistance [13], and decrease side effects by enabling reduced doses [1, 38]. While combination therapies confer benefits to cancer patients compared to single drug ones, personalizing drug combinations for individual patients is challenging. The effectiveness of drug combinations varies across patients due to differences in their genotypic and phenotypic variation [10, 27].

To facilitate drug combination discovery, preclinical studies rely on screening drug combinations in cancer cell lines experimentally. However, even for medium-sized drug panels, exhaustive experimentation with all possible combinations is impossible as the space of possible combinations increases exponentially with the number of drugs [11, 24, 30]. To aid the combination discovery efforts, though most are not personalized, many *in slico* methods have been proposed [4, 25]. Various mathematical modeling techniques [39], network-based models [7], and classical machine learning techniques have been applied to predict drug synergy [2, 17, 21, 33]. These models have been trained on various types of data, including drug chemical properties, cell line untreated and/or treated gene expression [34], drug-target interaction [2], and protein-protein interaction networks [7]. Yet, the prediction performance was limited due to limited training data.

Over the last decade, the research community focused on compiling large scale drug synergy datasets and organized several challenes to develop more advanced methods to exploit these resources [2, 35]. For instance, AstraZeneca and Sanger [35] released more than ten thousand experimentally tested drug combinations,

where cell viability was measured across various cell lines. Specifically, this DREAM challenge considered $11,759$ drug combination screenings that span 85 compounds and 85 cell lines. Similarly, NCI-ALMANAC study[14], evaluated the combination effects of 104 FDA-approved cancer drugs on 60 cell lines. This dataset contains nearly $300,000$ drug pair-cell line combinations.

With the availability of large-scale training data, we [19] and others [17, 22, 28, 31, 37] have developed deep learning based solutions to identify synergistic drug combinations. Preuer et al. presented DeepSynergy, which takes the chemical structures and cancer cell line gene expression data as input to a fully-connected neural network to predict synergy of drug pairings. MatchMaker [19] uses a multi-modal architecture and employs three fully-connected neural network modules to learn cell line-specific representations of drugs and predicts drug combination synergy scores. Finally, in a recent approach, DeepDDS [31] uses a graph convolutional neural network to predict combination synergy scores. This approach represents molecules as graphs, where atoms correspond to nodes and bonds correspond to edges.

While the deep learning models showed superior performance in predicting synergistic drug pairs, two critical challenges stand in the way of deploying deep learning models to personalize predictions for cancer patients. Cancer cell line expression data, available for many cancer cell lines, has been used to contextualize the synergy effects in these prior works. However, cell line expressions are only a proxy for the patient expression data; and do not capture the individual patient variation. The models require a large amount of training data to learn to patient synergy scores, but measurements on patient-derived samples are not available at large. To tackle patient customized synergy prediction problem, He et al. proposed a computational-experimental setup where they use *in silico* prioritization and *ex vivo* testing in patient-derived samples to identify customized synergistic combinations for individual cancer patients. Since the available patient response data is small, they build a random forest model to learn single-drug dose-response curves of patients and use this model to predict drug combinations. Unlike the above-mentioned deep learning models, the model does not benefit from the available large-scale synergy measurements obtained on cell lines. To the best of our knowledge, no model can make use of both the available large-scale cell line data and the limited patient data to achieve patient-level drug synergy prediction. In this work, we address this problem.

To tackle the patient-specific synergy prediction task, we propose a deep learning-based solution (Figure 1). In the first stage, we train the model using the large-scale single drug response and synergy scores of drug pairs measured on cell lines (Figure 1.1). The drug response prediction module enables us later to calibrate our model with the limited single drug measurements on patient-derived samples. In the second stage, we fine-tune the models using gene expression and drug response data measured on patient-derived *ex vivo* samples. To customize the models for each patient, we present two novel patient fine-tuning strategies that make use of the patient single drug response data. (i) we use the pre-trained model and then fine-tune the model with drug sensitivity labels of all training patients data (Figure 1.2.a); (ii) We use patient-specific fine-tuning, in which we fine-tune the pre-trained model for each patient separately (Figure 1.2.b). We then predict test patients' synergy data with the corresponding models (Figures 1.3.a and 1.3.b). We demonstrate that these strategies accurately predict synergy for the patients while maintaining the model's comparable performance on cell lines. Currently, there is no comparable study that uses deep learning models with measurements on patient data; therefore, this model represents an advance toward personalizing drug synergy predictions.

2 Methodology

For a drug pair $< i, j >$ and a cell line k, we aim to predict pair synergy scores $y_{i,j,k} \in \mathbb{R}$ along with each drug's sensitivities in cell line k. We binarize the individual compound sensitivity scores (see Section 2.1). Let $s_{i,k} \in \{0, 1\}$, and $s_{j,k} \in \{0, 1\}$ represent drug i and j's individual sensitivity classes, where 0 indicates resistant and 1 indicates sensitive. We represent each drug with its chemical features d_i and d_j, and cancer cell lines with gene expression profiles c_k. The model is trained jointly to predict the two drugs' individual sensitivity classes and pair synergy score by using drug and cell line feature vectors (d_i, d_j and c_k).

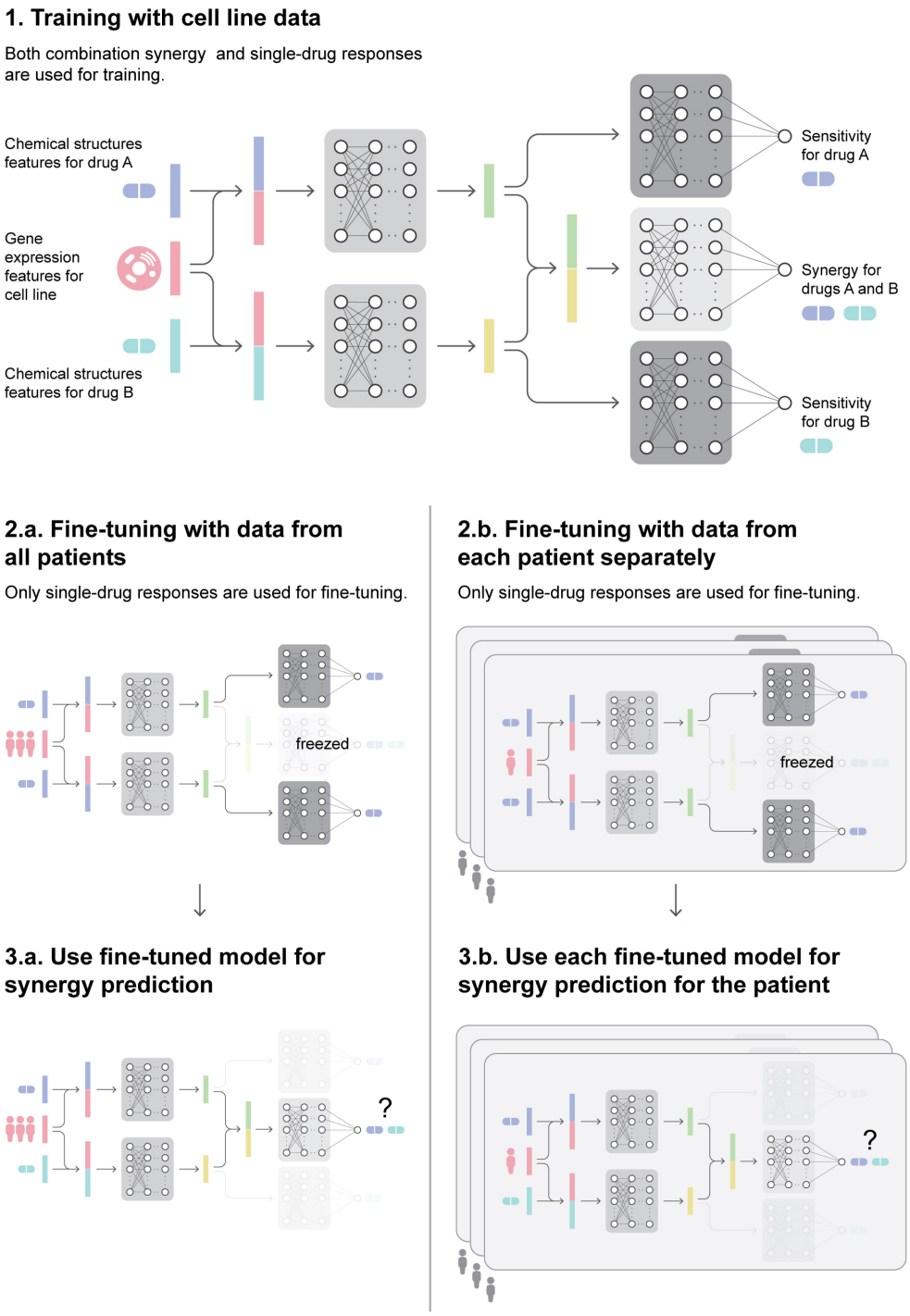

Fig. 1. The general pipeline for PDSP framework. 1) The architecture and the training procedure of the PDSP model. The chemical features of each drug are concatenated with gene expression features of the cancer cell line and input to drug encoder networks. They learn a representation of each drug separately conditioned on cancer cell line gene expressions. The synergy aggregations network takes both drugs' representations and applies fully connected layers to predict synergy scores. The drug decoder network takes conditioned drug representations and applies connected layers to predict the drug sensitivities. 2) Fine-tuning the trained PDSP model. The pre-trained PDSP model is used for fine-tuning the drug encoders by using patient single-drug responses and gene expressions for all patients (2.a) and per patient (2.b). 3) Synergy testing with respective fine-tuned models from the previous part.

2.1 Datasets

Cancer cell line drug response dataset: We use the Bliss synergy scores measured on cancer cell lines for synergy measurements and half-maximal inhibitory concentration (IC_{50}) values for sensitivity measurements. We obtain the data from the DrugComb [36] database through https://drugcomb.fimm.fi/ (version v1.5, downloaded on Feb 2022). We binarize IC_{50} values as sensitive and resistant. Following Chang et al., we designate cell lines as sensitive to a drug if the natural logarithm of its IC_{50} is less than -2 and resistant otherwise. Since DrugComb includes data from multiple sources, it sometimes contains replicates for the same drug pair. In these cases, we average Bliss scores across replicate measurements. Not all cancer cell line gene expression profiles were available at Iorio et al., so we filtered out cell lines for which gene expression profiles were not available. Our final dataset contains $330,103$ samples with $3,068$ drugs and 81 cancer cell lines.

Patient drug response and synergy dataset: We obtain patient synergy and sensitivity data from [12]. This work provides 654 sensitivity measurements on patient-derived samples covering 218 compounds on *ex vivo* and three patients. He et al. has also made drug combination experiments and reported 20 experimentally verified drug combinations results, of which 10 are synergistic, and the remaining 10 are antagonistic. These 20 combinations include 16 drugs for the three patients. We use the patient drug response data for finetuning the models and use the synergy measurements for evaluating the final model for personalized synergy prediction.

Model features: Our model relies on two sets of features, one set for representing drugs and the other set for cell line representation. We represent drugs with their chemical structure features, which we calculate using ChemoPy Python library [5] (more information provided in our earlier work [19]). Cell lines are represented with their untreated gene expression profiles [16], and RMA normalized profiles are obtained from https://www.cancerrxgene.org (downloaded on Dec 2019). Since we apply our model to patient expression profiles from He et al., we use available gene expression profiles from this study. To be able to apply PDSP on patient gene expression profiles, we take the intersection of $17,737$ genes in cancer cell lines and 634 genes in patient data. The intersection of these two gene sets contains 542 genes; we use expressions of these genes as gene expression features. The cancer cell line gene expressions and the patient gene expressions are measured under different conditions. We apply quantile normalization [3]. We input two drugs $d_i, d_j \in \mathbb{R}^{541}$ and a cell line $c_k \in \mathbb{R}^{542}$ to our model and predict $y_{i,j,k} \in \mathbb{R}$, $s_{i,k} \in \mathbb{R}$, and $s_{j,k} \in \mathbb{R}$ simultaneously.

2.2 Training Loss

We train the model with a loss function that combines the regression loss L_{reg} and the classification loss L_{cls} :

$$L = L_{reg} + \lambda * L_{cls} \tag{1}$$

where λ is a hyperparameter that adjusts the weight between regression and classification losses. We adopt *mean-squared-error* as the regression loss, and *weighted binary cross entropy* (wBCE) as the classification loss:

$$L_{reg} = \frac{1}{N} \sum_{i,j,k \in D} (y_{i,j,k} - \hat{y}_{i,j,k})^2 \tag{2}$$

$$\begin{aligned} L_{cls} &= wBCE(i,k) + wBCE(j,k) \\ wBCE_{t,k} &= -\frac{1}{N} \sum_{t,k \in D} \lambda_1 * s_{t,k} * log(\hat{s}_{t,k}) + \lambda_2 * (1 - s_{t,k}) * log(1 - \hat{s}_{t,k}) \end{aligned} \tag{3}$$

where N is the number of drug combinations in training dataset D. $y_{i,j,k}$ is the ground-truth Bliss score of drug combination $<i,j>$ on cell line k. $s_{t,k}$ is the ground-truth sensitivity class of drug t on cell line k. $\hat{y}_{i,j,k}$ and $\hat{s}_{t,k}$ are the corresponding predictions. λ_1 and λ_2 are the loss weights for sensitive and resistant drugs on cell line k.

2.3 The Model Architecture

PDSP model builds upon our previous work Matchmaker [19]. Figure 1 provides an overview of the PDSP architecture. We first represent cell lines and two drugs as a collection of vectors d_i, d_j, and c_k. With these vector representations, PDSP aggregates both cancer cell line and drugs' information and predicts drug sensitivities and combination synergy in two stages. In the first stage, PDSP encodes drug and cancer cell line information with Drug Encoder Networks (DEN). DENs take both drug and cancer cell line representations and learn conditional latent representations of the drugs. We force both DENs to share the same architecture and model weights to get order-invariant latent representations. In the second stage, PDSP uses these conditional encodings to predict sensitivity classes and combination synergy score. This second stage contains two Drug Decoder Networks (DDN) and one Synergy Aggregation Network (SAN). DDN takes encoded conditional latent representation and decodes it to the drug's sensitivity on a given cell line. Similarly to DENs, DDNs share the same architectures and model weights. While drug encoders learn representations of each drug, they do not provide information about their combinations in a cell line. Therefore, SAN first aggregates drugs' latent representation by concatenation and provides the combination information. Finally, an aggregated representation is used to predict drug combination synergy.

Drug Encoder Network (DEN): The PDSP model consists of two identical drug encoders, DEN_i and DEN_j. These encode the cancer cell line features (c_k) and two drugs' features (d_i and d_j) into a conditional latent space. The DEN model takes gene expression data and concatenates it with the features of each drug in the combination. This concatenated information is then passed through a series of fully-connected (FC) layers. Each FC layer is followed by a non-linear activation function (ReLU), except for the last FC layer and dropout layers. The formulation of DEN is as follows:

$$\begin{aligned} den_1(d,c) &= \text{Dropout}(\text{ReLU}(\text{FC}(\text{Concat}(d,c)))) \\ den_2(d,c) &= \text{Dropout}(\text{ReLU}(\text{FC}(den_1(d,c)))) \\ DEN(d,c) &= \text{Dropout}(\text{Linear}(\text{FC}(den_2(d,c)))) \end{aligned} \tag{4}$$

where d and c denotes the drug and cell line features, den_1 and den_2 are the intermediate layers. The DEN_i and DEN_j encoders are designed to have the same model parameters and weights, which allows for the same latent representations to be achieved for the two drugs (order-invariance). This shared structure enables the model to learn the relationships between the drugs and cancer cell lines consistently, regardless of the order in which the drugs are presented. The outputs of the encoders, $h_{i,k}$ and $h_{j,k} \in \mathbb{R}^m$, are multi-dimensional representations of the drugs in the conditional latent space. We use these drug encodings as inputs to downstream tasks, drug sensitivity, and drug combination prediction. DEN_i and DEN_j provide drugs encodings $h_{i,k}$ and $h_{j,k} \in \mathbb{R}^m$ as follows:

$$\begin{aligned} h_{i,k} &= \text{DEN}(\text{d}_\text{i}, \text{c}_\text{k}) \\ h_{j,k} &= \text{DEN}(\text{d}_\text{j}, \text{c}_\text{k}) \end{aligned} \tag{5}$$

Drug Decoder Network (DDN): Based on the latent drug encodings $h_{i,k}$ and $h_{j,k}$, we train two identical drug decoder networks DDN_i and DDN_j to output drug sensitivities $\hat{s}_{i,k}$ and $\hat{s}_{j,k}$. DDN uses a feed-forward neural network architecture with three FC layers. The first two FC layers apply and ReLU activation and dropout, respectively. Finally, a final FC layer with a linear activation function is used to obtain the drug sensitivity status. The DNN formulation is given below:

$$\begin{aligned} ddn_1(h) &= \text{Dropout}(\text{ReLU}(\text{FC}(h))) \\ ddn_2(h) &= \text{Dropout}(\text{ReLU}(\text{FC}(ddn_1(h)))) \\ DDN(s) &= \text{Linear}(\text{FC}(ddn_2(h))) \end{aligned} \tag{6}$$

where h corresponds to the conditional drug encodings ($h_{i,k}$ or $h_{j,k}$), ddn_1 and ddn_2 are the intermediate layers. Both DDNs use the same model weights, and sensitivity predictions $\hat{s}_{i,k}$ and $\hat{s}_{j,k}$ for drugs i and j on cancer cell line k is calculated as follows:

$$\begin{aligned} \hat{s}_{i,k} &= \mathrm{DDN}(\mathrm{h_{i,k}}) \\ \hat{s}_{j,k} &= \mathrm{DDN}(\mathrm{h_{j,k}}) \end{aligned} \tag{7}$$

Synergy Aggregation Network (SAN): Similarly to DDN, synergy aggregation network uses feed-forward neural network architecture. SAN takes latent drug representations $h_{i,k}$ and $h_{j,k}$ and predicts drug combination synergy $\hat{y}_{i,j,k}$. Since the inputs of SAN is a drug combination, it first concatenates drug features, then applies similar network architecture to DDN. We apply linear activation at the last layer of SAN since it outputs the Bliss synergy score. The SAN model is given below:

$$\begin{aligned} san_1(h_1, h_2) &= \mathrm{Dropout}(\mathrm{ReLU}(\mathrm{FC}(\mathrm{Concat}(h_1, h_2)))) \\ san_2(h_1, h_2) &= \mathrm{Dropout}(\mathrm{ReLU}(\mathrm{FC}(san_1(h_1, h_2)))) \\ SAN(h_1, h_2) &= \mathrm{Linear}(\mathrm{FC}(san_2(h_1, h_2))) \end{aligned} \tag{8}$$

where h_1 and h_2 denotes conditional drug encodings in the combination, san_1 and san_2 are the intermediate layers. With these formulations, we obtain predicted synergy score as $\hat{y}_{i,j,k} = SAN(h_{i,k}, h_{j,k})$.

2.4 Customizing for Patients

We initially train PDSP with drug combination data derived from cancer cell line experiments. In this pretrained model, we freeze SAN and fine-tune the model with patient single-drug responses. Fine-tuning provides us to calibrate DENs to the patient-derived sample data. We apply two patient fine-tuning strategies and asses their predictive performance.

We denote the initial trained model weights on cancer cell line as Φ_c, patients' gene expression profiles $c_{p_t} \in \mathbb{R}^{542}$ ($t \in \{1, 2, 3\}$), drugs' chemical features as $d_i \in \mathbb{R}^{541}$ and fine-tune Φ_c as follows:

Strategy 1: Fine-tune model with all patients: We use Φ_c as the pre-trained model weights for PDSP. We feed patients' gene expression profiles and drugs' chemical structure features into the model and run the PDSP training pipeline for single-drug responses. After fine-tuning, we use this new model weights $\Phi_{c,all}$ for synergy prediction of all patients. We refer to this model as $\mathrm{PDSP}^{\mathrm{st1}}$. This strategy is useful when we have a set of single-drug responses for multiple patients.

Strategy 2: Fine-tune model for each patient: In the second strategy, we fine-tune pre-trained weights Φ_c for each patient separately. This lead to n different models for each of the n patients: $\Phi_{c,t}$ ($t \in \{1, ...n\}$. We evaluate each patient's drug combination synergy score with the corresponding model's weights $\Phi_{c,t}$. We refer to the models generated using this strategy $\mathrm{PDSP}^{\mathrm{st2}}$.

3 Results

3.1 Experimental Setup

We follow Preuer et al. for performance benchmark and use *leave-drug-combination-out* approach. We divide the whole drug combination data into train, validation, and test splits by considering that a drug pair $<i, j>$ can only be involved in one of the splits. The ratios of train, validation and test splits are 60%, 20%, and 20%, respectively. We tune the model hyperparameters based on the performance of the validation set. We use these models to report performance on cancer cell line test data. We also use the best PDSP model weights Φ_c as the pre-trained model for patient data calibration as explained in 2.4.

3.2 Performance Benchmark

Regression Performance on Cancer Cell Lines We compare PDSP's performance to the state-of-the-art models: MatchMaker [19], DeepSynergy [28], TreeCombo [17] and DeepDDS [31]. We also use a standard machine learning algorithm, Random Forest, as the baseline model.We evaluate them based on MSE, Pearson, and Spearman rank correlation coefficients between the actual Bliss synergy scores and the predicted scores on the unseen cell line test set. Table 1 shows the model performances for these performance metrics. Among these methods, PDSP and MatchMaker achieve the top performances. Both models show comparable performances across all metrics. While MatchMaker shows slightly better performance on MSE (41.60 vs. 41.95), PDSP is better on the Spearman correlation coefficient (64% vs. 65%). They achieve the same performance in the Pearson correlation coefficient: 91%. These results show that adding drug response prediction nodes to enable patient calibration did not sacrifice on the synergy prediction performance; PDSP achieves on par synergy prediction performance on cell-lines with MatchMaker (i.e., no patient calibration). When we compare PDSP's performance with the remaining models, PDSP achieves $\sim$ 20% MSE and 2% to 7% correlation improvements over the next best model, DeepSynergy. DeepDDS, a model that uses graph neural networks to encode drug chemical structure, performs worse than the baseline method, Random Forest.

Table 1. Comparison of synery predictions of the models based on MSE, Pearson and Spearman correlation coefficients.

	MSE	**Pearson**	**Spearman**
PDSP	41.95	0.91	0.65
MatchMaker	41.60	0.91	0.64
DeepSynergy	50.65	0.89	0.59
TreeCombo	61.89	0.87	0.54
DeepDDS	177.33	0.54	0.26
Random Forest	165.72	0.60	0.31

We further assess the cell line specific performances of the models by measuring the regression performance metrics for each cell line separately. We provide the list of 10 cell lines for which our method performed the best. We also provide the number of instances that cell line is observed in training data in Supplementary Tables 1-3. We observe that the number of cases does not have a substantial effect on the performance.

Classification Performance on Cancer Cell Lines We also the classification performance of the synergy model. We binarize the actual Bliss scores: synergistic pairs have positive Bliss scores, and antagonistic pairs have negative Bliss scores. We perform binary classification to evaluate the predictive performance of the models and present the results in Table 2 using various metrics. Note that we do not perform any training for this task and use the discretized predictions and ground truth to measure classification performance.

PDSP and MatchMaker outperform other methods in regression performance. PDSP achieves better performances on AUPR (0.81 vs. 0.80) and Precision (0.70 vs. 0.69), MatchMaker is better on F1 (0.72 vs. 0.73) and Recall (0.75 vs. 0.76) metrics.

Table 2. Classification performance comparison of the models based on MSE, Pearson and Spearman correlation coefficients.

	AUC	**pAUC**	**AUPR**	**F1**	**Precision**	**Recall**	**Kappa**	**BACC**
PDSP	0.79	0.64	0.81	0.72	0.70	0.75	0.40	0.70
MatchMaker	0.79	0.64	0.80	0.73	0.69	0.76	0.40	0.70
DeepSynergy	0.76	0.62	0.78	0.70	0.68	0.73	0.35	0.68
DeepDDS	0.61	0.53	0.63	0.62	0.59	0.65	0.16	0.58
TreeCombo	0.73	0.60	0.75	0.68	0.67	0.79	0.32	0.66
Random Forest	0.63	0.55	0.66	0.54	0.63	0.48	0.17	0.59

Patient Personalized Classification Performance While we compare the synergy prediction performance of PDSP with others on cancer cell lines and show above that it is on par with the state-of-the-art methods, its ultimate goal is to perform personalized predictions for cancer patients. Cancer cell line measurements provide large scale training data, but they fail to capture patient variation as cancer is a complex and heterogeneous disease. We have very limited patient-level ground truth data to train a complex machine learning model. For this reason, we employ transfer learning approach in the design of PDSP and use the single drug responses of patients to fine-tune the model as explained in Section 2.4. We compare the classification performances of PDSP (uncalibrated), MatchMaker, DeepSynergy, TreeCombo, DeepDDS and Random Forest models with two patient-calibrated PDSP^{st1} and PDSP^{st2} models. Figure 2 shows the number of *True Positives* (TPs), *True Negatives* (TNs), *False Positives* (FPs) and *False Negatives* (FNs).

Figure 2 shows that uncalibrated models are not good at detecting antagonistic drug combinations. For example, Random Forest method predicts all drug pair-patient trios as synergistic. Similarly, MatchMaker and TreeCombo predict 17 and 18 (out of 20) drug pairs as synergistic, respectively. They mainly act as a one-class-predictor, which leads to a large number of FPs and low number of TNs. These results lead to low precision values (47% and 56%, respectively) for these models. Compared to uncalibrated models, PDSP^{st1} and PDSP^{st2} achieve better precision (63% and 83%) and separation between synergistic and antagonistic drug combinations. Specifically, the patient-calibrated PDSP^{st1} model decreases the number of FPs to 3 and increases the number of TNs to 7 compared to PDSP. This shows that fine-tuning with patient gene expression improves precision (63% vs. 43%) and accuracy (60% vs. 40%). PDSP^{st2} model performs even better with 9 TNs at the cost of 1 FP. This indicates that it is nearly perfect at detecting antagonistic drug combinations while preserving the performance at synergistic drug combinations. Thus, using fine-tuning and using per-patient calibration (Strategy 2) in particular, yields the best-personalized drug synergy prediction performance.

Pat. ID	Compound 1	Compound 2	Random Forest	TreeCombo	DeepSynergy	MatchMaker	DeepDDS	PDSP	PDSP st1	PDSP st2
1263	FIMM003747-003	FIMM003787-001	+	+	-	+	+	+	+	-
1263	FIMM000618-004	FIMM109450-001	+	+	+	+	+	-	-	-
1263	FIMM100381-001	FIMM003747-003	+	+	+	-	+	-	-	-
1263	FIMM109441-001	FIMM003747-003	+	+	+	+	+	+	+	+
1263	FIMM100388-001	FIMM115484-001	+	+	-	+	-	-	-	-
1263	FIMM001244-002	FIMM115484-001	+	+	+	+	-	+	+	-
1409	FIMM109442-001	FIMM003735-001	+	+	+	+	-	+	-	+
1409	FIMM115616-002	FIMM003735-001	+	+	-	+	-	-	+	+
1409	FIMM109442-001	FIMM115616-002	+	+	+	+	-	+	+	+
1409	FIMM000296-002	FIMM000618-004	+	+	+	+	-	+	-	-
1409	FIMM115616-002	FIMM100346-001	+	+	+	+	-	-	-	+
1409	FIMM100381-001	FIMM003747-003	Text	+	+	-	+	-	-	-
1409	FIMM109441-001	FIMM003747-003	+	-	+	+	+	+	+	+
1508	FIMM003790-001	FIMM003767-001	+	+	-	+	+	+	-	-
1508	FIMM000296-002	FIMM000618-004	+	+	-	+	+	+	-	-
1508	FIMM000618-004	FIMM109450-001	+	+	+	+	+	+	-	-
1508	FIMM000296-002	FIMM109450-001	+	+	+	-	+	+	-	-
1508	FIMM100381-001	FIMM003747-003	+	-	+	+	+	+	-	-
1508	FIMM109441-001	FIMM003747-003	+	+	+	+	+	+	+	-
1508	FIMM001244-002	FIMM115484-001	+	+	+	+	-	+	+	-
		TRUE +	10	10	7	9	4	6	5	5
		TRUE -	0	2	2	2	2	2	7	9
		FALSE +	10	8	8	8	8	8	3	1
		FALSE -	0	0	3	1	6	4	5	5
		ACCURACY	11/20	12/20	9/20	11/20	6/20	8/20	12/20	14/20

Fig. 2. The classification results of patient drug combination synergy data. All models are trained with DrugComb data with cell line gene expressions. PDSP^{st1} and PDSP^{st2} correspond to patient-calibrated models that are fine-tuned based on Strategy 1 and Strategy 2 respectively.

4 Discussion

This study is the first attempt to perform in-silico personalized drug synergy prediction for cancer patients that uses large-scale gene expression data. With limited patient-level data, we opted for a transfer learning based approach to first train a complex model using large scale cell-line based drug-pair perturbation datasets and used patients' drug response profiles to fine-tune and personalize the model. We show favorable results compared to the state-of-the-art models that do not allow incorporating personal features.

The results indicate that fine-tuning a cell line pre-trained model with a small number of single-drug responses enables PDSP to identify synergistic and antagonistic drug pairs for specific patients better than other models. While the uncalibrated PDSP model's accuracy is 40% when tested for patients, fine-tuned PDSP^{st1} and PDSP^{st2} models' accuracies are 60% and 70%, respectively. We observe that uncalibrated models tend to make more synergistic predictions, and they cannot detect antagonistic drug pairs. This problem is alleviated by fine-tuning.

This study uses the drug's chemical structure and cell line specific gene expression profiles as the feature set. As indicated in [19], these features are selected as they are widely available for most drugs and provide a large training set to train a complex model like ours. If available on a large scale, our model can always be supplemented with other sources of information such as drug-disease or drug-target interactions and network-based methods [7]. Similarly, it is straightforward to adapt the model to use other synergy scores such as Loewe, ZIP and HSA.

We hypothesize that personal single-drug responses can be indirectly used to fine-tune the PDSP model toward better personal drug synergy prediction. We modify the MatchMaker architecture to use this information. We show that, indeed it benefits the drug synergy prediction task in patients. However, as personal drug synergy labels become widely available, we foresee that transfer learning using actual synergy labels will result in even better performance. This is because models will be able to incorporate directly relevant labels, and the model architecture will be less complex with fewer parameters to optimize.

Bibliography

[1] Al-Lazikani, B., Banerji, U., Workman, P.: Combinatorial drug therapy for cancer in the post-genomic era. Nature biotechnology **30**(7), 679 (2012)

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

[3] Bolstad, B.M., Irizarry, R.A., Åstrand, M., Speed, T.P.: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics **19**(2), 185–193 (2003)

[4] Bulusu, K.C., Guha, R., Mason, D.J., Lewis, R.P., Muratov, E., Motamedi, Y.K., Cokol, M., Bender, A.: Modelling of compound combination effects and applications to efficacy and toxicity: state-of-the-art, challenges and perspectives. Drug discovery today **21**(2), 225–238 (2016)

[5] Cao, D.S., Xu, Q.S., Hu, Q.N., Liang, Y.Z.: ChemoPy: freely available python package for computational biology and chemoinformatics. Bioinformatics **29**(8), 1092–1094 (03 2013), ISSN 1367-4803, https://doi.org/10.1093/bioinformatics/btt105, URL https://doi.org/10.1093/bioinformatics/btt105

[6] Chang, Y., Park, H., Yang, H.J., Lee, S., Lee, K.Y., Kim, T.S., Jung, J., Shin, J.M.: Cancer drug response profile scan (cdrscan): a deep learning model that predicts drug effectiveness from cancer genomic signature. Scientific reports **8**(1), 1–11 (2018)

[7] Cheng, F., Kovács, I.A., Barabási, A.L.: Network-based prediction of drug combinations. Nature communications **10**(1), 1–11 (2019)

[8] Clercq, E.D.: The design of drugs for hiv and hcv. Nature reviews Drug discovery **6**(12), 1001–1018 (2007)

[9] Csermely, P., Korcsmáros, T., Kiss, H.J., London, G., Nussinov, R.: Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review. Pharmacology & therapeutics **138**(3), 333–408 (2013)

[10] Day, D., Siu, L.L.: Approaches to modernize the combination drug development paradigm. Genome medicine **8**(1), 115 (2016)

[11] He, L., Kulesskiy, E., Saarela, J., Turunen, L., Wennerberg, K., Aittokallio, T., Tang, J.: Methods for high-throughput drug combination screening and synergy scoring. In: Cancer systems biology, pp. 351–398, Springer (2018)

[12] He, L., Tang, J., Andersson, E.I., Timonen, S., Koschmieder, S., Wennerberg, K., Mustjoki, S., Aittokallio, T.: Patient-customized drug combination prediction and testing for t-cell prolymphocytic leukemia patients. Cancer research **78**(9), 2407–2418 (2018)

[13] Hill, J.A., Ammar, R., Torti, D., Nislow, C., Cowen, L.E.: Genetic and genomic architecture of the evolution of resistance to antifungal drug combinations. PLoS genetics **9**(4), e1003390 (2013)

[14] Holbeck, S.L., Camalier, R., Crowell, J.A., Govindharajulu, J.P., Hollingshead, M., Anderson, L.W., Polley, E., Rubinstein, L., Srivastava, A., Wilsker, D., et al.: The national cancer institute almanac: a comprehensive screening resource for the detection of anticancer drug pairs with enhanced therapeutic activity. Cancer research **77**(13), 3564–3576 (2017)

[15] Home, P.D., Pocock, S.J., Beck-Nielsen, H., Curtis, P.S., Gomis, R., Hanefeld, M., Jones, N.P., Komajda, M., McMurray, J.J., Team, R.S., et al.: Rosiglitazone evaluated for cardiovascular outcomes in oral agent combination therapy for type 2 diabetes (record): a multicentre, randomised, open-label trial. The Lancet **373**(9681), 2125–2135 (2009)

[16] Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., et al.: A landscape of pharmacogenomic interactions in cancer. Cell **166**(3), 740–754 (2016)

[17] Janizek, J.D., Celik, S., Lee, S.I.: Explainable machine learning prediction of synergistic drug combinations for precision cancer medicine. bioRxiv p. 331769 (2018)

[18] Kopetz, S., Grothey, A., Yaeger, R., Van Cutsem, E., Desai, J., Yoshino, T., Wasan, H., Ciardiello, F., Loupakis, F., Hong, Y.S., et al.: Encorafenib, binimetinib, and cetuximab in braf v600e–mutated colorectal cancer. New England Journal of Medicine **381**(17), 1632–1643 (2019)

[19] Kuru, H.I., Tastan, O., Cicek, E.: Matchmaker: a deep learning framework for drug synergy prediction. IEEE/ACM Transactions on Computational Biology and Bioinformatics (2021)
[20] Lee, J., Yesilkanal, A.E., Wynne, J.P., Frankenberger, C., Liu, J., Yan, J., Elbaz, M., Rabe, D.C., Rustandy, F.D., Tiwari, P., et al.: Effective breast cancer combination therapy targeting bach1 and mitochondrial metabolism. Nature **568**(7751), 254–258 (2019)
[21] Li, P., Huang, C., Fu, Y., Wang, J., Wu, Z., Ru, J., Zheng, C., Guo, Z., Chen, X., Zhou, W., et al.: Large-scale exploration and analysis of drug combinations. Bioinformatics **31**(12), 2007–2016 (2015)
[22] Liu, Q., Xie, L.: Transynergy: mechanism-driven interpretable deep neural network for the synergistic prediction and pathway deconvolution of drug combinations. PLoS computational biology **17**(2), e1008653 (2021)
[23] Liu, T., Yacoub, R., Taliaferro-Smith, L.D., Sun, S.Y., Graham, T.R., Dolan, R., Lobo, C., Tighiouart, M., Yang, L., Adams, A., et al.: Combinatorial effects of lapatinib and rapamycin in triple-negative breast cancer cellscombined treatment in triple-negative breast cells. Molecular cancer therapeutics **10**(8), 1460–1469 (2011)
[24] Macarron, R., Banks, M.N., Bojanic, D., Burns, D.J., Cirovic, D.A., Garyantes, T., Green, D.V., Hertzberg, R.P., Janzen, W.P., Paslay, J.W., et al.: Impact of high-throughput screening in biomedical research. Nature reviews Drug discovery **10**(3), 188–195 (2011)
[25] Menden, M.P., Wang, D., Mason, M.J., Szalai, B., Bulusu, K.C., Guan, Y., Yu, T., Kang, J., Jeon, M., Wolfinger, R., et al.: Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen. Nature communications **10**(1), 1–17 (2019)
[26] Palmer, A.C., Sorger, P.K.: Combination cancer therapy can confer benefit via patient-to-patient variability without drug additivity or synergy. Cell **171**(7), 1678–1691 (2017)
[27] Pang, K., Wan, Y.W., Choi, W.T., Donehower, L.A., Sun, J., Pant, D., Liu, Z.: Combinatorial therapy discovery using mixed integer linear programming. Bioinformatics **30**(10), 1456–1463 (2014)
[28] Preuer, K., Lewis, R.P., Hochreiter, S., Bender, A., Bulusu, K.C., Klambauer, G.: Deepsynergy: predicting anti-cancer drug synergy with deep learning. Bioinformatics **34**(9), 1538–1546 (2017)
[29] Sun, Y., Sheng, Z., Ma, C., Tang, K., Zhu, R., Wu, Z., Shen, R., Feng, J., Wu, D., Huang, D., et al.: Combining genomic and network characteristics for extended capability in predicting synergistic drugs for cancer. Nature communications **6**(1), 1–10 (2015)
[30] Torres, N.P., Lee, A.Y., Giaever, G., Nislow, C., Brown, G.W.: A high-throughput yeast assay identifies synergistic drug combinations. Assay and drug development technologies **11**(5), 299–307 (2013)
[31] Wang, J., Liu, X., Shen, S., Deng, L., Liu, H.: Deepdds: deep graph neural network with attention mechanism to predict synergistic drug combinations. Briefings in Bioinformatics **23**(1), bbab390 (2022)
[32] Wang, R., Chan, J.F.W., Wang, S., Li, H., Zhao, J., Ip, T.K.Y., Zuo, Z., Yuen, K.Y., Yuan, S., Sun, H.: Orally administered bismuth drug together with n-acetyl cysteine as a broad-spectrum anti-coronavirus cocktail therapy. Chemical science **13**(8), 2238–2248 (2022)
[33] Wildenhain, J., Spitzer, M., Dolma, S., Jarvik, N., White, R., Roy, M., Griffiths, E., Bellows, D.S., Wright, G.D., Tyers, M.: Prediction of synergism from chemical-genetic interactions by machine learning. Cell Systems **1**(6), 383–395 (2015)
[34] Yang, J., Tang, H., Li, Y., Zhong, R., Wang, T., Wong, S., Xiao, G., Xie, Y.: Digre: Drug-induced genomic residual effect model for successful prediction of multidrug effects. CPT: pharmacometrics & systems pharmacology **4**(2), 91–97 (2015)
[35] Yu, T.: Astrazeneca-sanger drug combination prediction dream challenge (2015)
[36] Zagidullin, B., Aldahdooh, J., Zheng, S., Wang, W., Wang, Y., Saad, J., Malyutina, A., Jafari, M., Tanoli, Z., Pessia, A., et al.: Drugcomb: an integrative cancer drug combination data portal. Nucleic acids research **47**(W1), W43–W51 (2019)
[37] Zhang, T., Zhang, L., Payne, P.R., Li, F.: Synergistic drug combination prediction by integrating multiomics data in deep learning models. Translational bioinformatics for therapeutic development pp. 223–238 (2021)
[38] Zhao, S., Nishimura, T., Chen, Y., Azeloglu, E.U., Gottesman, O., Giannarelli, C., Zafar, M.U., Benard, L., Badimon, J.J., Hajjar, R.J., et al.: Systems pharmacology of adverse event mitigation by drug combinations. Science translational medicine **5**(206), 206ra140–206ra140 (2013)
[39] Zhao, X.M., Iskar, M., Zeller, G., Kuhn, M., Van Noort, V., Bork, P.: Prediction of drug combinations by integrating molecular and pharmacological data. PLoS computational biology **7**(12) (2011)