Predicting drug synergy for PTCL patient-derived xenografts with models trained on cell lines by transfer learning

This project includes development of a deep learning model, DeepPTCL, to generating transfer predictors of drug synergy from cell lines to PDX Tumors. DeepPTCL achieves superior prediction performance in DrugComb benchmark dataset during the training phase(AUC = 0.82). Our contribution can be stated as follows: 1) We propose a novel DL layer for pair scoring and apply it to drug synergy predictions 2) We conduct an transfer learning to highlight the potential drug combinations on PTCL patient PDX tumors.

Chengqi Xu

Created on December 12 | Last edited on December 29

Dec 12, 2023, Inghirami lab & Elemento lab

DeepPTCL accurately predict drug synergy in cell lines

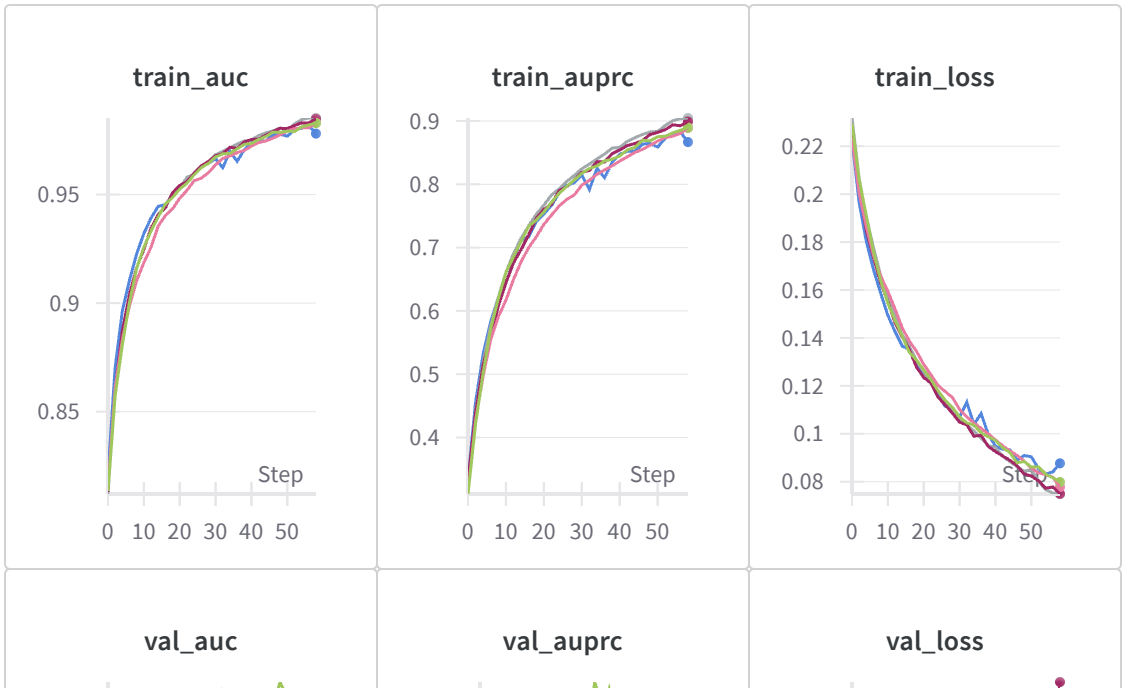
TCGA cell lines have been the workhorse of cancer research, producing massive amounts of drug response data. Unfortunately, translating response biomarkers derived from these datasets to PDX tumors has proven to be particularly challenging. To address this challenge, we introduce **DeepPTCL** , a novel deep learning model designed to predict drug synergy on cell lines from The Cancer Genome Atlas. DeepPTCL achieves state-of performances on drug

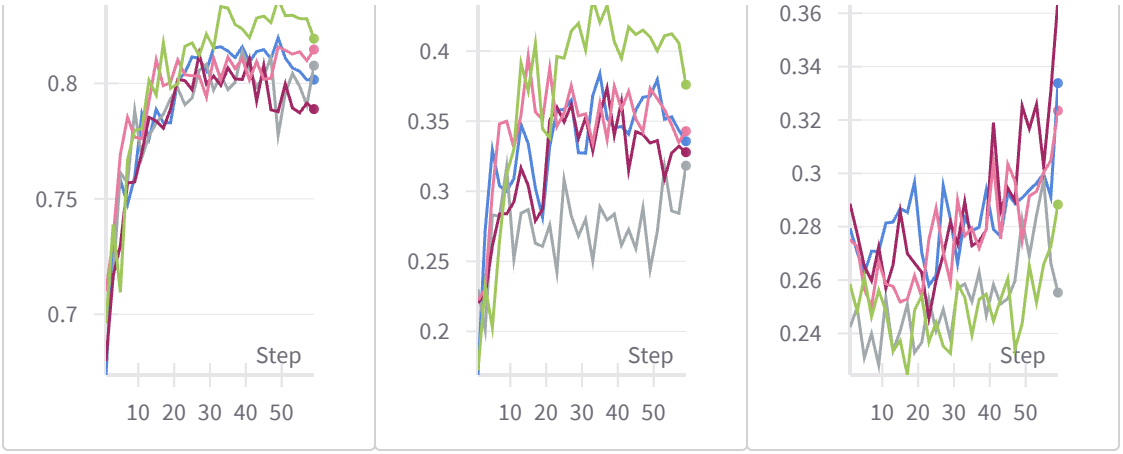
synergy prediction challenges in the largest public synergism dataset, DrugComb. We further apply it to PTCL patient tumors to guide combination therapy in a personalized medicine approach.

A. Dataset details: Drugs represented as SMILES were converted using RDKit into a PyG graph, with atoms represented as nodes and bonds represented as edges. The features of the cell line are gathered from the Genomics of Drug Sensitivity in Cancer. From the normalized expression levels of 17,737 genes, we select 908 landmark genes. To develop the model, we used DrugComb (the largest database of high-throughput combination screening data) as a training dataset with 2,174 drugs on 164 cell lines, altogether a total of 16,3816 drug-drug-cell triplets.

B. Training of DeepPTCL: We trained the model a 5-fold cross-validation (Each colored line represents a run). We designed two challenges tasks for training to improve model generalizability. In the leave-drug-out setup, we excluding drugs seeing from training from the test set. For the leave-combination-out setup, the drug pairs from training are removed from the test set, although the same individual drugs might appear in both dataset. The performance of the model is evaluated using AUROC and AUPRC metrics.

The interactive panel shows superior performance of DeepCTL in benchmarks.





Guiding Drug Combination in the in-vivo setting of PDX tumors

TCGA cell lines and PDX tumor data are in a consensus space

We next wished to move beyond cell lines to predict and interpret drug responses in the in vivo setting of patient-derived xenograft models. The genotypes of each PDX had also been established, which were provided to DeepPTCL to make predictions to each combination therapy. Of note previous published models while promising in performance, but do not fully take into account the fundamental differences between preclinical models and human tumors or only model these differences as a technical batch effect. To address this challenge, we developed a mathematical approach to successfully gene expression patterns shared between preclinical 164 TCGA cell lines and 207 PDX tumors in a consensus space.

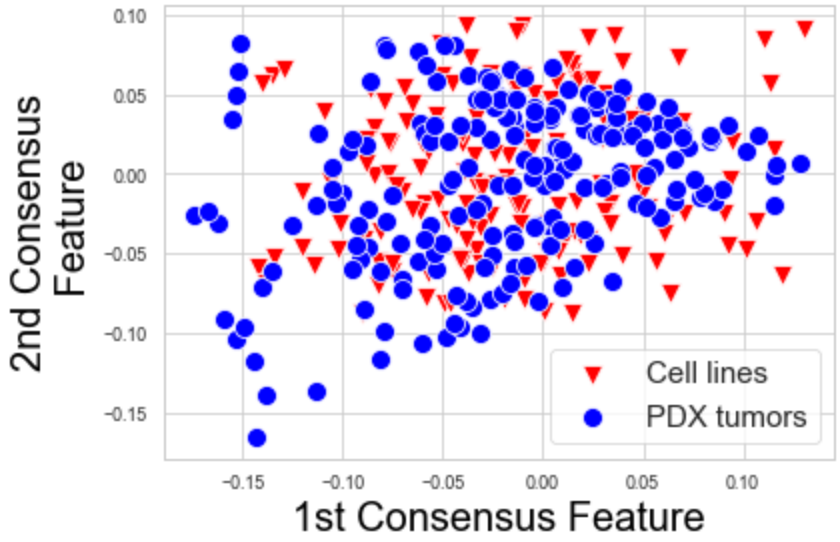


Fig 1. Our approach compares genomic signals contained in the source (e.g., TCGA cell lines) and target (e.g., PDX tumors) datasets and outputs a consensus space. By using the Gaussian similarity function, various types of nonlinear relationships can be used to as predictors transferred from cell Lines to PDXs.

DeepPTCL finds drug combinations for PDX tumors to 8 small molecules that have been used and/or offered to PTCL patients

For each PDX tumor, measured gene express values in DeepPTCL landmark genes were used as input to DeepPTCL to predict the drug synergy to 8 drugs (Irinotecan, Romedepsin, Duvelisib, Pralatrexate, AZD-4573, Cedulatinib, Azacytidine, Crizotinib) combined with secondary drugs . We design two in sillico screenings described as follows. 1) The pairwise combinations within 8 drugs yield 8*7*207=11,592 drug-drug-cells combinations. 2) (Future direction) The secondary drugs can be those 2,147 oncology-focused compounds that DeepPTCL had previously seen and had combination data available in the DrugComb training dataset, which will yield 8*2,147*207=3,555,432 drug-drug-cells combination searching space with secondary drugs from diverse target classes.

Top 6 predicted synergy combinations are prioritized with an dual PI3K inhibtor Duvelisib

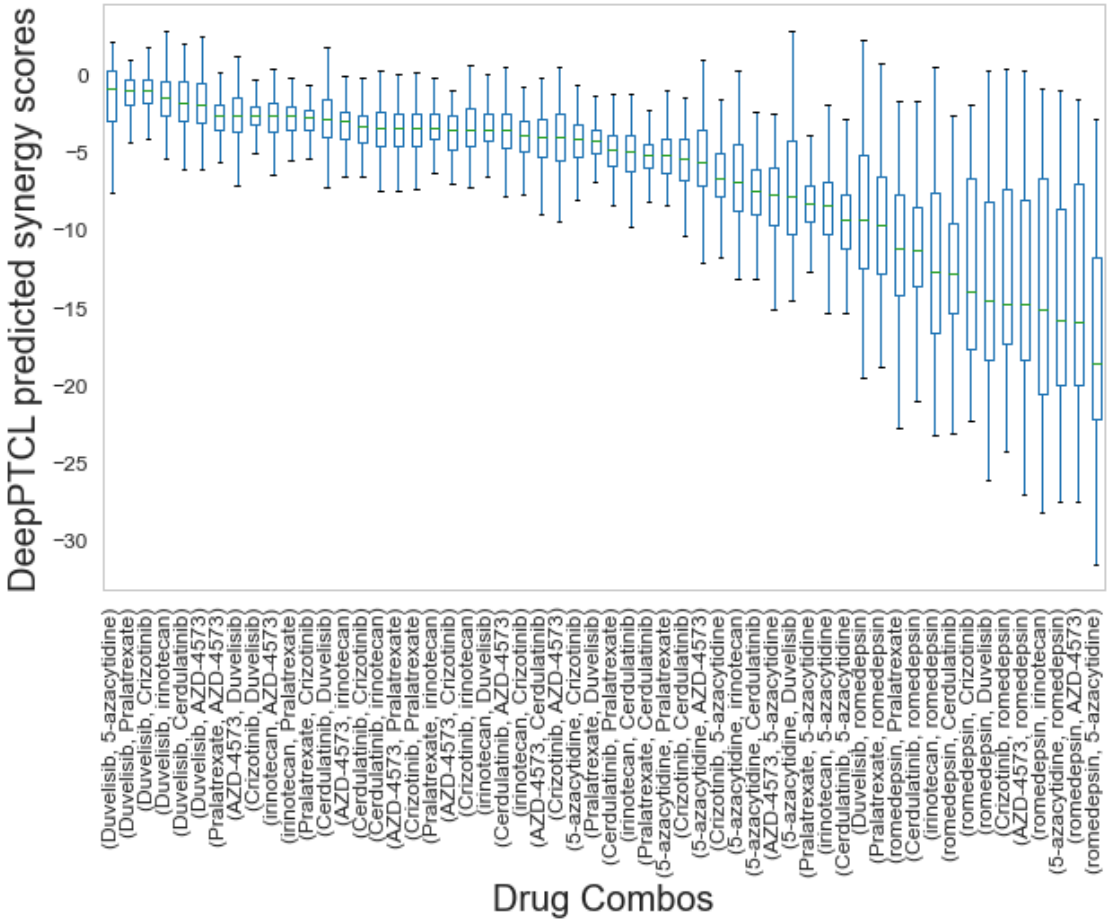


Fig 2. The distribution of DeepPTCL predicted synergism across pairwise combinations within 8 drugs. Boxplot are sorted by the median value of predicted synergy scores across 207 PDX tumors for each pairwise combination sampled within 8 drugs. Top 6 drug combinations are seen to be synergistic with an anchor drug Duvelisib, where the combining drugs prioritized by our models are 5-azacytidine, pralatrexate, crizotinib, irinotecan, cerdulatinib, and AZD4573 respectively.

Patient-specific combinations show high overall synergy in Duvelisib-containing drug combinations and wide interpatient variability

The predicted synergies of the patient-specific combination are higher in Naive, BIA-ACL and AITL compared to activated or cALCL types.

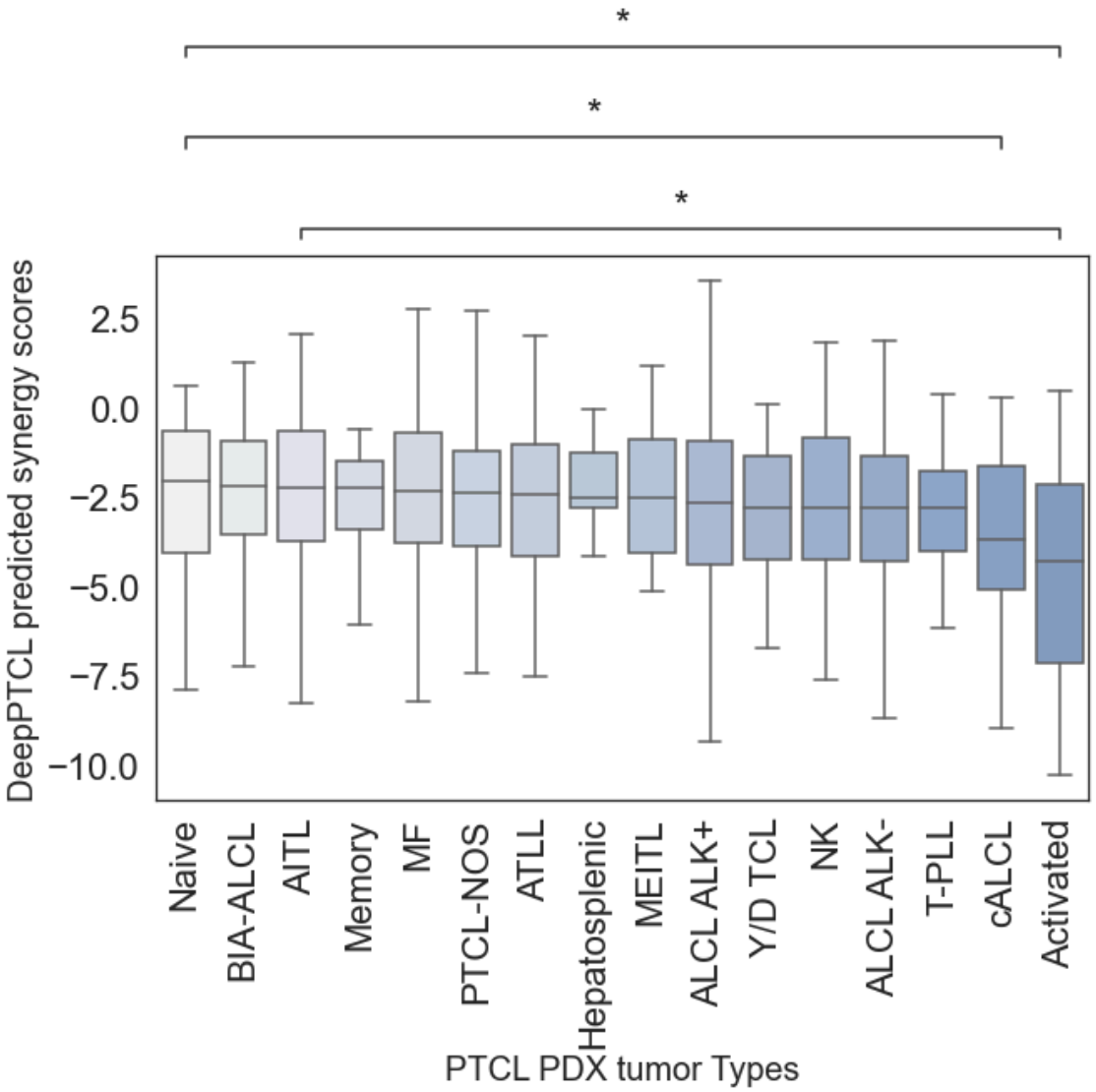


Fig 3. Patient-specific combinations predicted uniquely for each patient sample type. p-value annotation legend: *: 1.00e-02 < p <= 5.00e-02. Mann-Whitney-Wilcoxon wo-sided test was used to find the u-statistic and p-value..AITL vs. Activated: P_val:1.017e-02 U_stat=1.202e+04; Naive vs. cALCL:P_val:4.926e-02 U_stat=2.361e+03; Naive vs. Activated:P_val:2.936e-02 U_stat=1.833e+03

Created with ❤️ on Weights & Biases.

https://wandb.ai/chloexyz123/giorgio/reports/Predicting-drug-synergy-for-PTCL-patient-derived-xenografts-with-models-trained-on-cell-lines-by-transfer-learning--Vmlldzo2MjQ3NTUx