Predicting Synergistic Drug Combinations for COVID with Biological Bottleneck Models

Wengong Jin **Regina Barzilay** **Tommi Jaakkola**
CSAIL, Masssachusetts Institute of Technology
{wengong,regina,tommi}@csail.mit.edu

Abstract

Drug combinations play an important role in therapeutics due to its better efficacy and reduced toxicity. Recent approaches have applied machine learning to identify synergistic combinations for cancer, but they are not applicable to new diseases with limited combination data. Given that drug synergy is closely tied to biological targets, we propose a *biological bottleneck* model that jointly learns drug-target interaction and synergy. The model consists of two parts: a drug-target interaction and target-disease association module. This design enables the model to *explain* how a biological target affects drug synergy. By utilizing additional biological information, our model achieves 0.78 test AUC in drug synergy prediction using only 90 COVID drug combinations for training.

1 Introduction

Combination therapies have shown to be more effective than single drugs in multiple diseases such as HIV and tuberculosis [25, 28]. Synergistic combinations can improve both potency and efficacy, either achieving stronger therapeutic effects and/or decreasing dosage thereby reducing side-effects. In the times of current pandemic, finding a successful combination of approved molecules have an additional benefit over designing a de-novo molecule: time to clinical adoption. Approved drugs are typically commercially available and have well studied safety profiles. Taken in aggregate, these considerations motivate us to explore combination therapies for SARS-CoV-2 antivirals.

Since exploring the space of combinations via high-throughput screening is prohibitively expensive as it involves combinatorial search, in-silico screening based on machine learning is an appealing alternative. In fact, a number of such methods have been reported in the literature [22, 26]. These techniques have been shown effective when the model was provided with large amounts of training data capturing synergy of various combinations. Unfortunately, this requirement prevents us from utilizing these techniques for many diseases where such data is not available. Therefore, it is crucial to reduce data dependence to make combination algorithm applicable in multiple therapeutic contexts.

In this paper, we present a novel algorithm for finding combinations that achieves this goal. Our main hypothesis is that by explicitly modeling interaction between compounds and the biological targets, we can significantly decrease dependence on combination training data. The proposed *biological bottleneck model* has two components. The first component models drug-target interactions (DTI) predicting which targets are inhibited by a compound. It is trained on individual compounds since DTI information is readily available for multiple targets across multiple diseases. Our second component focuses on modeling target-disease association. It is a simple linear function which enables the model to explain how much a biological targets affects synergistic activity.

We develop our model using single agent and drug combination data from various sources. It incorporates known COVID biological targets [9] and their corresponding drug-target activity collected from ChEMBL [7]. With only 90 COVID drug combinations for training, our model achieves 0.78 test

Machine Learning for Molecules Workshop at NeurIPS 2020. https://ml4molecules.github.io

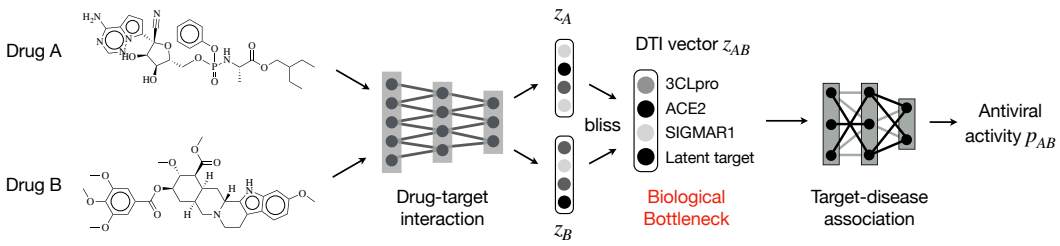

Figure 1: Our biological bottleneck model is composed of two modules: a drug-target interaction (DTI) and target-disease association module. The antiviral effect of a combination is predicted on their DTI vector $\boldsymbol{z}_{AB}$, which is computed from the DTI vectors $\boldsymbol{z}_A, \boldsymbol{z}_B$ of each individual drug. The DTI vector characterizes the drug-target interaction profile of a given drug.

AUC on the SARS-CoV-2 combination screen from Bobrowski et al. [2]. Moreover, incorporating known COVID targets yields 10% relative increase in test accuracy. We have experimentally tested our model predictions in the NCATS facilities and reported the discovered novel synergistic drug combinations in `https://arxiv.org/pdf/2011.04651.pdf`.

2 Biological Bottleneck Models

In this section, we describe our model architecture for drug combinations. A drug combination is called synergistic if its antiviral effect is greater than the sum of the individual effects. Drug synergy arises from various types of drug interaction. For example, two drugs can be synergistic when they interact with different sets of biological targets or pathways. Indeed, most of the anti-HIV drug combinations, such as Dolutegravir and Lamivudine, are drugs with different mechanisms of actions (i.e., interacting with different biological targets). To account for this inductive bias, it is crucial to model the interaction between drugs and biological targets in our model architecture.

Motivated by these observations, we propose to decompose our model into two parts: a drug-target interaction (DTI) module ϕ and a target-disease association module f. The DTI module predicts the biological targets activated by a given compound. The target-disease association module learns how a biological target is related to the disease. The vocabulary of biological targets are chosen by experts in advance. To introduce our method, we first describe how these two modules are used to predict antiviral activity of single compounds and then extend it to drug combinations.

2.1 Forward pass of single drugs

We represent a drug as a graph $\mathcal{G}$, whose nodes and edges represent atoms and bonds. To predict the antiviral effect of a single drug A, our model needs to accomplish two tasks: 1) predict its interaction with biological targets $\mathcal{V} = \{t_1, \cdots, t_m\}$; 2) learn the relevance of each target t_i to the disease.

Drug-target interaction We parametrize the DTI module ϕ as a graph convolutional network (GCN) [5, 8]. The GCN translates a molecular graph $\mathcal{G}_A$ into a continuous vector through directed message passing operations [27], which associate hidden vectors $\boldsymbol{h}_v$ with each node v and updates these vectors by passing messages $\boldsymbol{h}_{uv}$ over edges (u, v). The output of ϕ is a vector $\boldsymbol{z}_A$ representing the biological targets activated by drug A:

$$\boldsymbol{z}_A = \sigma\big(\mathrm{MLP}\big(\sum\nolimits_{v \in \mathcal{G}_A} \boldsymbol{h}_v\big)\big) \qquad \{\boldsymbol{h}_v\} = \mathrm{GCN}(\mathcal{G}_A) \tag{1}$$

where $\sigma(\cdot)$ is a sigmoid function and MLP is a two-layer feed-forward network. Each element $\boldsymbol{z}_{A,k}$ represents the probability of drug A inhibits target t_k. Each target t_k is associated with a drug-target interaction dataset $\mathcal{D}_k = \{(X_i, y_i)\}$, where $y_i = 1$ if a drug X_i is interacts with target t_k. We will train this module on the DTI dataset of all biological targets in the vocabulary.

Target-disease association We parametrize the target-disease association module f as a simple linear layer $(\boldsymbol{w}, \boldsymbol{b})$ due to its interpretability. As shown in Figure 1, our model predicts the antiviral activity of a drug A as:

$$p_A = f(\boldsymbol{z}_A) = \sigma(\boldsymbol{w}^\top \boldsymbol{z}_A + \boldsymbol{b}) \tag{2}$$

2.2 Forward pass of drug combinations

Synergy are often quantified under Bliss synergy score [1]. Suppose the individual antiviral effect of drugs A and B are p_A, p_B. The expected effect of combination (A, B) is given as $e_{AB} = p_A + p_B - p_A p_B$. A drug combination (A, B) is synergistic if its observed effect $p_{AB} > e_{AB}$. Following this definition, we introduce a new Bliss layer to predict the synergistic effect of a drug combination (A, B). Given two drugs and their predicted DTI vectors $\boldsymbol{z}_A, \boldsymbol{z}_B$, the Bliss layer computes the DTI vector $\boldsymbol{z}_{AB}$ as

$$\boldsymbol{z}_{AB} = \boldsymbol{z}_A + \boldsymbol{z}_B - \boldsymbol{z}_A \odot \boldsymbol{z}_B \tag{3}$$

where $\odot$ stands for element-wise multiplication. With this aggregation function, a drug combination will benefit most from *complementary* targets. If only one drug is active to target t_i (e.g., $\boldsymbol{z}_{A,i} = 1, \boldsymbol{z}_{B,i} = 0$), the combination (A, B) is still active to t_i ($\boldsymbol{z}_{AB,i} = 1$). In other words, the set of active targets for (A, B) is the union of active targets of the two drugs. Given a drug combination (A, B), our model predicts its antiviral activity as:

$$p_{AB} = f(\boldsymbol{z}_{AB}) = \sigma(\boldsymbol{w}^\top \boldsymbol{z}_{AB} + \boldsymbol{b}) \tag{4}$$

Following Bliss independence model, we predict the synergy score of a combination as $p_{AB} - e_{AB}$, where $e_{AB} = p_A + p_B - p_A p_B$. Intuitively, a combination is more likely to be synergistic if they have complementary targets with high target-disease association score.

2.3 Learning latent targets

In order to predict synergy, it is important to incorporate all the relevant biological targets into our model. However, this is challenging for two reasons: First, most biological targets do not have drug-target interaction data and thus cannot be incorporated in our model. Second, current biological understanding of a disease may be incomplete. For instance, Riva et al. [23] reported around 50 new biological targets related to SARS-CoV-2 antiviral activity, but they are not reported in the previous work by Gordon et al. [9].

To this end, we propose to include additional *latent targets* in the bottleneck layer that are learned indirectly from single-agent and combination data. Specifically, we expand the dimension of $\boldsymbol{z}_A$ to be greater than the total number of considered targets in $\mathcal{V}$. The first m entries in $\boldsymbol{z}_A$ corresponds to the real biological targets and the other entries are latent targets. As we will show in the experiments, it is possible for us to interpret new biological targets related to given diseases.

2.4 Training

Our training loss $\mathcal{L} = \lambda_{\text{DTI}} \ell_{\text{DTI}} + \lambda_S \ell_S + \ell_C$ consists of three components. The drug-target interaction loss ℓ_{DTI} enforces the predicted DTI vector $\boldsymbol{z}_A$ to be biologically meaningful. The DTI module ϕ is trained to minimize the loss on DTI training set $\mathcal{D}_k = \{(X_i, y_i)\}$: $\ell_{\text{DTI}} = \sum_k \sum_{(X_i, y_i) \in \mathcal{D}_k} \ell(\boldsymbol{z}_{X_i,k}, y_i)$ The single-agent prediction loss is computed on a single-agent training set $\mathcal{D}_S = \{(X_1, y_1), \cdots, (X_n, y_n)\}$, which contains molecules and their labeled antiviral activity (active/inactive). Both modules ϕ, f are trained to minimize $\ell_S = \sum_{(X_i, y_i) \in \mathcal{D}_S} \ell(f(\phi(X_i)), y_i)$ The combination loss is computed on a dataset $\mathcal{D}_C = \{(A_i, B_i, y_i)\}$ of drug combinations and their labeled synergy. $y_i = 1$ means (A_i, B_i) is synergistic and $y_i = 0$ additive or antagonistic. We train both modules ϕ, f to minimize $\ell_C = \sum_{(A_i, B_i, y_i) \in \mathcal{D}_C} \ell(p_{A_i B_i} - e_{A_i B_i}, y_i)$

Multi-disease training Since COVID is a new disease, its drug combination data is very limited. To address the low-resource challenge, we utilize additional drug synergy data from other viral diseases such as HIV. Specifically, we augment the model with HIV biological targets as well as HIV single-agent and drug combination data. The DTI module ϕ now outputs a DTI vector $\boldsymbol{z}_A^d$ for each disease d. ϕ is shared across two diseases and trained to learn drug-target interaction for all diseases. Since each disease operates on different targets, we create a target-disease association module f^D for each disease. Let $\ell_{\text{DTI}}^d, \ell_S^d, \ell_C^d$ be the losses for each disease $d \in \{\text{COVID, HIV}\}$. Our final training loss becomes $\mathcal{L}_{\text{multi}} = \sum_d \lambda_1 \ell_{\text{DTI}}^d + \lambda_2 \ell_S^d + \ell_C^d$

3 Experiments

SARS-CoV-2 Data The training data for SARS-CoV-2 comes from the following assays:

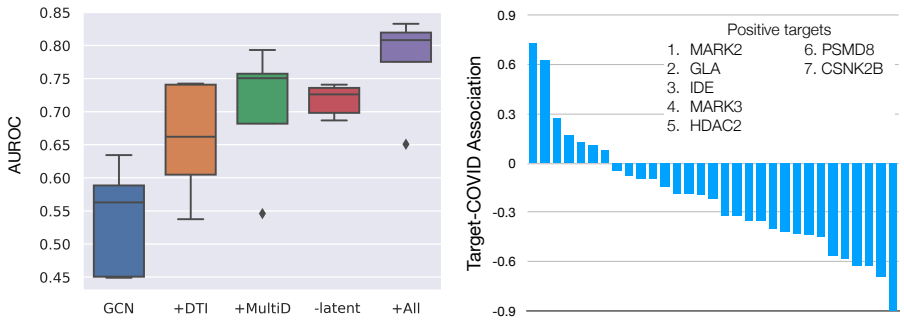

Figure 2: *Left*: Results on SARS-CoV-2 combination test set. Our model (+all) outperforms all other baselines. *Right*: Seven targets that positively contributes COVID drug synergy.

- *Drug-target Interaction*: Our target set includes viral proteases (3CLpro, PLpro) [15, 4], ACE2 for viral entry [11, 19] and 31 human proteins that are physically associated with SARS-CoV-2 viral proteins [9] and have sufficient amount of drug-target interaction data in the ChEMBL database.
- *Single-agent Activity*: We use the NCATS CPE assay in VeroE6 cells [17], which contains around 10K compounds and 320 hits with EC50 $\leq 10\mu$M.
- *Drug Combination*: NCATS performed two combination assays in VeroE6 cells, which contain 160 two-drug combinations [2, 18]. Riva et al. [23] also analyzed synergy between Remdesivir and 20 active compounds identified from their high-throughput screen.

HIV Data The training data for HIV comes from the following assays:

- *Drug-target Interaction*: Existing anti-HIV drugs mainly target viral proteins (HIV-1 protease, integrase and reverse transcriptase) or host proteins involved in viral entry (CCR5, CXCR4 and CD4). We compiled DTI data for these six targets from ChEMBL.
- *Single-agent Activity*: NCI conducted an anti-HIV assay [20] with 35K compounds, among which 309 compounds are active (EC50 $\leq 1\mu$M).
- *Drug Combination*: Tan et al. [25] conducted high-throughput screen for HIV drug combinations. The dataset contains 114 two-drug combinations.

Evaluation Protocol Since our goal is to predict synergy against SARS-CoV-2, our validation and test set only consist of SARS-CoV-2 combinations. All the drug-target interaction, single-drug activity and HIV data are used for training only. Our validation set contains 20 combinations from Riva et al. [23] and test set contains 72 combinations from Bobrowski et al. [2]. The training set contains 90 SARS-CoV-2 combinations from [18], where we remove combinations that appear in both the training and test set.

Hyperparameters For DTI module ϕ, we adopt default hyperparameters from Yang et al. [27], with hidden dimension 300 and three message passing iterations. We set the dimension of DTI vector $|\boldsymbol{z}| = 100$, with 42 real biological targets (SARS-CoV-2 and HIV) and 58 latent targets, so that the number of real and latent targets are roughly equal. We set $\lambda_1 = 10, \lambda_2 = 0.1$ for our final model.

Results To show the effectiveness of different components, we compare with the following baselines:

- A GCN trained only on SARS-CoV-2 single-agent and combination data ($\lambda_{\text{DTI}} = 0, |\boldsymbol{z}| = 100$).
- +DTI: A GCN trained only on SARS-CoV-2 single-agent, combination as well as drug-target interaction data ($\lambda_{\text{DTI}} = 10, |\boldsymbol{z}| = 100$).
- +MultiD: A GCN trained on both SARS-CoV-2 and HIV data (single-agent + combination), but without drug-target interaction data ($\lambda_{\text{DTI}} = 0, |\boldsymbol{z}| = 100$).
- +All,-latent: A GCN trained on both SARS-CoV-2 and HIV data (single-agent + combination + drug-target interaction), but the latent targets are removed ($\lambda_{\text{DTI}} = 10, |\boldsymbol{z}| = 42$).
- +All: A GCN trained on both SARS-CoV-2 and HIV data (single-agent + combination + drug-target interaction) ($\lambda_{\text{DTI}} = 10, |\boldsymbol{z}| = 100$).

Our results are shown in Figure 2. As expected, the GCN baseline performs poorly, with 0.537 ± 0.075 AUC. Adding drug-target interaction data (+DTI) improves the AUC to 0.658 ± 0.079. Adding HIV data (+MultiD) improves the AUC to 0.706 ± 0.088. Our final model, trained with both HIV and

drug-target interaction (+All), achieves the best AUC of 0.777 ± 0.066. This validates the advantage of adding drug-target interaction data and multi-disease training. Note that if we remove the latent targets (+All,-latent), the performance decreases to 0.718 ± 0.021. This also shows the importance of using latent targets to complement missing biological information.

References

[1] Chester I Bliss. The toxicity of poisons applied jointly 1. *Annals of applied biology*, 26(3): 585–615, 1939.

[2] Tesia Bobrowski, Lu Chen, Rich T Eastman, Zina Itkin, Paul Shinn, Catherine Chen, Hui Guo, Wei Zheng, Sam Michael, Anton Simeonov, et al. Discovery of synergistic and antagonistic drug combinations against sars-cov-2 in vitro. *BioRxiv*, 2020.

[3] Feixiong Cheng, István A Kovács, and Albert-László Barabási. Network-based prediction of drug combinations. *Nature communications*, 10(1):1–11, 2019.

[4] Meredith Davis-Gardner. Sars-cov-2 plpro inhibitor assay. 2020. `https://reframedb.org/assays/A00461`.

[5] David K Duvenaud, Dougal Maclaurin, Jorge Iparraguirre, Rafael Bombarell, Timothy Hirzel, Alán Aspuru-Guzik, and Ryan P Adams. Convolutional networks on graphs for learning molecular fingerprints. In *Advances in neural information processing systems*, pages 2224–2232, 2015.

[6] Bernhard Ellinger, Denisa Bojkova, Andrea Zaliani, Jindrich Cinatl, Carsten Claussen, Sandra Westhaus, Jeanette Reinshagen, Maria Kuzikov, Markus Wolf, Gerd Geisslinger, et al. Identification of inhibitors of sars-cov-2 in-vitro cellular toxicity in human (caco-2) cells using a large scale drug repurposing collection. 2020.

[7] Anna Gaulton, Anne Hersey, Michał Nowotka, A Patricia Bento, Jon Chambers, David Mendez, Prudence Mutowo, Francis Atkinson, Louisa J Bellis, Elena Cibrián-Uhalte, et al. The chembl database in 2017. *Nucleic acids research*, 45(D1):D945–D954, 2017.

[8] Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. Neural message passing for quantum chemistry. *arXiv preprint arXiv:1704.01212*, 2017.

[9] David E Gordon, Gwendolyn M Jang, Mehdi Bouhaddou, Jiewei Xu, Kirsten Obernier, Kris M White, Matthew J O'Meara, Veronica V Rezelj, Jeffrey Z Guo, Danielle L Swaney, et al. A sars-cov-2 protein interaction map reveals targets for drug repurposing. *Nature*, pages 1–13, 2020.

[10] Deisy Morselli Gysi, Ítalo Do Valle, Marinka Zitnik, Asher Ameli, Xiao Gan, Onur Varol, Helia Sanchez, Rebecca Marlene Baron, Dina Ghiassian, Joseph Loscalzo, et al. Network medicine framework for identifying drug repurposing opportunities for covid-19. *arXiv preprint arXiv:2004.07229*, 2020.

[11] Markus Hoffmann, Hannah Kleine-Weber, Simon Schroeder, Nadine Krüger, Tanja Herrler, Sandra Erichsen, Tobias S Schiergens, Georg Herrler, Nai-Huei Wu, Andreas Nitsche, et al. Sars-cov-2 cell entry depends on ace2 and tmprss2 and is blocked by a clinically proven protease inhibitor. *Cell*, 181(2):271–280, 2020.

[12] Susan L Holbeck, Richard Camalier, James A Crowell, Jeevan Prasaad Govindharajulu, Melinda Hollingshead, Lawrence W Anderson, Eric Polley, Larry Rubinstein, Apurva Srivastava, Deborah Wilsker, 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.

[13] Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Adaptive invariance for molecule property prediction. *arXiv preprint arXiv:2005.03004*, 2020.

[14] S Loewe. Die quantitativen probleme der pharmakologie. *Ergebnisse der Physiologie*, 27(1): 47–187, 1928.

[15] National Center for Advancing Translational Sciences (NCATS). Sars-cov-2 3cl protease enzymatic activity. 2020. https://opendata.ncats.nih.gov/covid19/assay?aid=9.

[16] National Center for Advancing Translational Sciences (NCATS). Ace2 enzymatic activity. 2020. https://opendata.ncats.nih.gov/covid19/assay?aid=6.

[17] National Center for Advancing Translational Sciences (NCATS). Sars-cov-2 cytopathic effect (cpe) screening. 2020. https://opendata.ncats.nih.gov/covid19/assay?aid=14.

[18] National Center for Advancing Translational Sciences (NCATS). In vitro sars-cov-2 evaluation of drug combinations for exploring chemical biology and potential therapeutic use. 2020. https://tripod.nih.gov/matrix-client/?project=2983.

[19] National Center for Advancing Translational Sciences (NCATS). Sars-cov-2 spike-ace2 protein-protein interaction (alphalisa). 2020. https://opendata.ncats.nih.gov/covid19/assay?aid=1.

[20] National Cancer Institute (NCI). Aids antiviral screen data. 2004. https://wiki.nci.nih.gov/display/NCIDTPdata/AIDS+Antiviral+Screen+Data.

[21] Jennifer O'Neil, Yair Benita, Igor Feldman, Melissa Chenard, Brian Roberts, Yaping Liu, Jing Li, Astrid Kral, Serguei Lejnine, Andrey Loboda, et al. An unbiased oncology compound screen to identify novel combination strategies. *Molecular cancer therapeutics*, 15(6):1155–1162, 2016.

[22] Kristina Preuer, Richard PI Lewis, Sepp Hochreiter, Andreas Bender, Krishna C Bulusu, and Günter Klambauer. Deepsynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics*, 34(9):1538–1546, 2018.

[23] Laura Riva, Shuofeng Yuan, Xin Yin, Laura Martin-Sancho, Naoko Matsunaga, Lars Pache, Sebastian Burgstaller-Muehlbacher, Paul D De Jesus, Peter Teriete, Mitchell V Hull, et al. Discovery of sars-cov-2 antiviral drugs through large-scale compound repurposing. *Nature*, pages 1–11, 2020.

[24] Pavel Sidorov, Stefan Naulaerts, Jérémy Ariey-Bonnet, Eddy Pasquier, and Pedro Ballester. Predicting synergism of cancer drug combinations using nci-almanac data. *Frontiers in chemistry*, 7:509, 2019.

[25] Xu Tan, Long Hu, Lovelace J Luquette, Geng Gao, Yifang Liu, Hongjing Qu, Ruibin Xi, Zhi John Lu, Peter J Park, and Stephen J Elledge. Systematic identification of synergistic drug pairs targeting hiv. *Nature biotechnology*, 30(11):1125–1130, 2012.

[26] Fangfang Xia, Maulik Shukla, Thomas Brettin, Cristina Garcia-Cardona, Judith Cohn, Jonathan E Allen, Sergei Maslov, Susan L Holbeck, James H Doroshow, Yvonne A Evrard, et al. Predicting tumor cell line response to drug pairs with deep learning. *BMC bioinformatics*, 19(18):71–79, 2018.

[27] Kevin Yang, Kyle Swanson, Wengong Jin, Connor Coley, Philipp Eiden, Hua Gao, Angel Guzman-Perez, Timothy Hopper, Brian Kelley, Miriam Mathea, et al. Analyzing learned molecular representations for property prediction. *Journal of chemical information and modeling*, 59 (8):3370–3388, 2019.

[28] Kaan Yilancioglu and Murat Cokol. Design of high-order antibiotic combinations against m. tuberculosis by ranking and exclusion. *Scientific reports*, 9(1):1–11, 2019.

[29] Yadi Zhou, Yuan Hou, Jiayu Shen, Yin Huang, William Martin, and Feixiong Cheng. Network-based drug repurposing for novel coronavirus 2019-ncov/sars-cov-2. *Cell discovery*, 6(1):1–18, 2020.

A Related Work

Existing approach on drug synergy prediction can be roughly divided into two categories:

- *Supervised learning*: In this approach, the model is trained on combination data generated from high-throughput screens. For example, Preuer et al. [22] have trained a deep neural network on a large-scale oncology screen [21] (23K training examples) to predict anti-cancer drug synergy. Xia et al. [26] and Sidorov et al. [24] have trained deep neural networks to predict anti-cancer drug synergy on a larger dataset compiled by NCI-ALMANAC [12], which contains around 300K training examples across 40 different cell lines.
- *Biological networks*: Another category of drug synergy models builds on biological networks, assuming that drugs with complementary mechanism of actions are more likely to be synergistic. In particular, Cheng et al. [3], Zhou et al. [29] proposed to model synergy using distance measures on drug-target interaction and protein-protein interaction networks.

The major challenge of supervised approaches is the lack of combination data. For many diseases such as SARS-CoV-2 and tuberculosis, the size of drug combination data is very limited (less than 200) [2, 28]. Deep learning methods are prone to over-fitting in this low-resource scenario. Moreover, the number of possible pair-wise combinations grows quadratically with the number of drugs. In fact, the current largest combination screen for cancer [12] only covers around 100 different drugs. This significantly limits the ability of trained models to generalize to new drugs outside of the training set. On the other hand, while network-driven methods have a wider coverage over the chemical space, they cannot make predictions on new compounds outside of the biological network (i.e., no edge between this compound and the targets) since the model is not parametric.

We propose a new method that combines the merit of both approaches while addressing their limitations. As drug interaction is often characterized by their biological targets, we train our model to predict both drug-target interaction and drug synergy. This enables us to make predictions on new compounds even if their drug-target interaction is unknown. This also addresses the data scarcity challenge since there are large amounts of drug-target interaction publicly available.

B Biological Targets

For SARS-CoV-2 infection, we consider three types of biological targets in our target vocabulary $\mathcal{V} = \{t_1, \cdots, t_m\}$:

- *Viral proteases*: Replication of SARS-CoV-2 virus requires the processing of two polyproteins by two virally encoded proteases: chymotrypsin-like protease (3CLpro) and papain-like protease (PLpro). Inhibitors that block either protease could inhibit viral replication. We have compiled 3CLpro enzymatic activity [15] and PLpro inhibition [4] data made public by NCATS and ReframeDB.
- *Viral entry proteins*: SARS-CoV-2 cell entry depends on angiotensin converting enzyme 2 (ACE2) [11]. Inhibiting ACE2 enzyme or the interaction between SARS-CoV-2 and ACE2 could block viral entry. To this end, we utilize ACE2 enzymatic activity [16] and Spike-ACE2 protein-protein interaction [19] from NCATS.
- *Host proteins*: Gordon et al. [9] identified 335 human proteins physically associated with SARS-CoV-2 viral proteins. Inhibitors for these proteins may also hinder viral replication. Among these proteins, we selected 31 proteins that have sufficient amount of drug-target interaction data in the ChEMBL database (i.e., both positive and negative interactions).