DEEMD: Drug Efficacy Estimation against SARS-CoV-2 based on cell Morphology with Deep multiple instance learning

by

Mohammadsadegh Saberian

B.Sc., Sharif University of Technology, 2018

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
School of Computing Science
Faculty of Applied Sciences

SIMON FRASER UNIVERSITY
Spring 2021

Declaration of Committee

Name:	**Mohammadsadegh Saberian**
Degree:	**Master of Science**
Thesis title:	**DEEMD: Drug Efficacy Estimation against SARS-CoV-2 based on cell Morphology with Deep multiple instance learning**
Committee:	**Chair:** Ke Li Assistant Professor Computing Science
	Maxwell Libbrecht Co-Supervisor Assistant Professor, Computing Science
	Ghassan Hamarneh Co-Supervisor Professor, Computing Science
	Ivan R. Nabi Committee member Professor, Cellular and Physiological Sciences University of British Columbia
	Martin Ester Examiner Professor, Computing Science

Abstract

Background: Drug repurposing can accelerate the identification of effective compounds for clinical use against SARS-CoV-2, with the advantage of pre-existing clinical safety data and an established supply chain. RNA viruses such as SARS-CoV-2 manipulate cellular pathways and induce reorganization of subcellular structures to support their life cycle. These morphological changes can be quantified using bioimaging techniques.

Methods: In this work, we developed DEEMD: a computational pipeline using deep neural network models within a multiple instance learning (MIL) framework, to identify putative treatments effective against SARS-CoV-2 based on morphological analysis of the publicly available RxRx19a dataset, This dataset consists of fluorescence microscopy images of SARS-CoV-2 non-infected cells and infected cells, with and without drug treatment. DEEMD first extracts discriminative morphological features to generate cell morphological profiles from the non-infected and infected cells. These morphological profiles are then used in a statistical model to estimate the applied treatment efficacy on infected cells based on similarities to non-infected cells.

Results: DEEMD is capable of localizing infected cells via weak supervision without any expensive pixel-level annotations. DEEMD identifies known SARS-CoV-2 inhibitors, such as *Remdesivir* and *Aloxistatin*, supporting the validity of our approach.

Conclusions: DEEMD is scalable to process and screen thousands of treatments in parallel and can be applied to other emerging viruses and data sets to rapidly identify candidate antiviral treatments in the future.

Keywords: Drug Repurposing; Deep Multiple Instance Learning; Morphological Analysis; SARS-CoV-2

Dedication

To whom I tremendously love...

Acknowledgements

First of all, I would like to thank my co-supervisors Prof. Ghassan Hamarneh and Prof. Maxwell W. Libbrecht. I was lucky to have two great mentors in my graduate school who guided me through my journey. Both Max and Ghassan provided me with financial support for my studies and made sure I am on the right track during my research.

Needless to say, this project will not succeed without the help and contribution of my colleagues and collaborators; so I would like to thank those who worked with me in this project. I would like to acknowledge Dr. Andrea D. Olmstead (Department of Microbiology and Immunology, University of British Columbia), Kathleen P. Moriarty (School of Computing Science, Simon Fraser University), Dr. François Jean (Department of Microbiology and Immunology, University of British Columbia), and Dr. Ivan R. Nabi (Life Sciences Institute, University of British Columbia) for their contribution to this research.

I would like to thank Prof. Martin Ester for serving as my examiner, Dr. Ivan R. Nabi (Life Sciences Institute, University of British Columbia) for being on my supervisory committee, and Prof. Li for serving as the chair for my thesis defence.

This study was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Alliance COVID-19 grant. I would like to thank Natural Sciences and Engineering Research Council of Canada Collaborative Research and Training Experience (NSERC-CREATE) program for supporting my studies.

Last but not least, I would like to thank my wonderful wife who unconditionally supported and helped me throughout my studies; and my family for their unparalleled love throughout my life.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The COVID-19 global pandemic has urged the research community to focus their resources towards studying SARS-CoV-2 and discovering or identifying potential therapeutics. To date, despite intense efforts, very few treatment options are available for those suffering from COVID-19 [67]. Drug repurposing is an attempt to identify existing clinically approved treatments with established pharmacological and safety profiles that could be rapidly redirected towards clinical treatment of novel diseases such as COVID-19 [78, 109] (and references within). The antiviral activity of candidate compounds can be tested using cell-based systems of viral infection. The detection of the viral infection is achieved using molecular tools such as antibodies directed at virus encoded proteins. In the case of newly emerging viruses such as SARS-CoV-2, access to such molecular tools may represent an important limiting step to rapidly developing cell-based assays to discover novel antiviral agents. Alternatively, since human pathogenic viruses such as SARS-CoV-2 manipulate cellular pathways to reorganize the host cell morphology to support their life cycles [58, 18], developing a computational method to perform quantitative analysis of virus-induced cell morphology provides a unique approach to discover candidate antiviral molecules without the use of viral biomarkers [8]. Virus infected cells can be treated with thousands of compounds at different concentrations followed by staining of cellular structures with fluorescent dyes that can be imaged using high-content screening fluorescence microscopes [74, 88]. Morphological features can be extracted from images of infected and uninfected cells and then applied to images of infected drug-treated cells to predict antiviral efficacy based on cellular morphology. Thus, quantitative morphological analysis of cells as a computational method for drug repurposing allows for accelerated parallel screening of multiple therapeutics [63].

It is worth mentioning that an early version of this work was accepted and presented at Machine Learning in Computational Biology (MLCB) 2020[1]. An extended version is under review in IEEE Transactions on Medical Imaging.

[1] `https://sites.google.com/cs.washington.edu/mlcb2020/`

The rest of this thesis is structured as follows: we first review the literature on drug repurposing, high-throughput cell imaging and how it can be used to scale drug repurposing studies, along with the computational multiple instance learning framework in Chapter 2. Chapter 3 presents DEEMD and its components and discusses them in detail. We applied DEEMD to a dataset of SARS-CoV-2 infected florescence microscopy images, as explained in Chapter 4, and identified a list of 18 potential treatments effective for COVID-19. We discuss our findings along with DEEMD's limitations in Chapter 5. Finally, the this thesis is concluded in Chapter 6.

Chapter 2

Background and Literature Review

2.1 Coronavirus

Named after its spike-like peplomers covering its surface, coronavirus has been known to human since 1960s and were mostly associated with mild respiratory disease [64, 52, 45, 14, 12, 86, 91]. In the past 20 years, three major outbreaks have been related to this family of viruses. During the first one which occurred in 2002–2003, over 8000 people from 26 countries were infected with severe acute respiratory syndrome coronavirus (SARS-CoV) which had a reported 10% case fatality rate. The Middle East respiratory syndrome coronavirus (MERS-CoV) outbreak started around April 2012 in Middle East, Africa and South Asia. Until November 2019, around 2500 labratory-confirmed cases of MERS-CoV have been reported involving 27 countries. It is estimated that MERS-CoV has a mortality rate of approximately 34% based only on the laboratory-confirmed cases. However, researches believe this mortality rate is an overestimation, since mild cases of MERS-CoV would be missed by the monitoring systems since little was known about it [3].

The most recent outbreak, presumably originating from Wuhan in China, started from a Huanan wholesale seafood market in December 2019 and quickly became a global pandemic. To this date, more than 128 million positive Coronavirus cases have been reported globally. Unfortunately, near 2.8 million people have lost their lives to it. Coronavirus disease (COVID-19), caused by SARS-CoV-2, has a variety of symptoms ranging from flu or gastrointestinal irritation to acute respiratory distress syndrome, heart arrhythmias, stroke, and death [34, 95, 1, 70, 99, 56, 6, 49]. Due to its high infection rate, diagnosis and therapy for COVID-19 has been the major focus of researches for the past few months. Approval of a novel therapeutic for COVID-19 can take several years, however, there has been major progress on development and roll out of a vaccine. As of today, at least seven different vaccines across three platforms have been rolled out in countries. At the same time, more than 200 additional vaccine candidates are in development, of which more than 60 are in clinical development [69, 54, 59, 21].

2.2 Drug Repurposing and High-throughput Imaging

Drug repurposing is an active research area in the pharmaceutical industry in which existing drugs are used in alternative applications for which they were originally designed. Drug repurposing has been shown to be an effective approach in many cases [78]. For instance, Fingolimod was originally developed to prolong organ graft survival. But later through drug repurposing, it was found to be effective against multiple sclerosis pathogenic immune responses [9]. Drug repurposing studies develop and apply various approaches to test their repurposing hypotheses. These approaches include computational methods, such as genetic association or molecular docking, and experimental ones, *e.g.* phenotypic screening or binding assays. However, deploying large-scale binding cell-based assays for large-scale systematic analysis is expensive and slow to the point that it would not be practical. On the other hand, high-throughput imaging assays can be used to screen numerous compounds for morphological cell profiling and eventually their antiviral effectiveness with high efficiency. Morphological cell profiling combined with recent advancements in computer vision models can be leveraged to extract the morphological changes induced by each compound in the sample population. These morphological features can lead to a better understanding of the compounds interacting within the host cell, its molecular targets, and its pathways. Image-based morphological analysis has been effectively utilized in small molecule profiling [8, 97], identifying the mechanism of action [50, 20, 89], and drug repurposing [84, 63].

Recently, drug repurposing studies focused on COVID-19 to find FDA-approved clinically understandable therapeutic compounds against SARS-CoV-2 [115, 51] (and references within). Among multiple methodologies, high-throughput screenings have enabled image-based morphological analysis of cells infected with SARS-CoV-2 [63, 38, 22, 75]. Mirabelli *et al.* [63] used Huh-7 cell line samples infected with SARS-CoV-2 and applied a library of 1425 FDA-approved compounds at different concentrations for identifying compounds with antiviral activity. The samples were later stained with anti-nucleocapsid protein, SARS-CoV-2 antibody, and dyes specific to cell organelles for imaging. They used CellProfiler software [62] to extract morphological features from the infected cells, using anti-nucleocapsid protein as indicator for regions of interest, and trained a random forest model to predict efficacy scores for each treatment. These scores were used to select efficacious compounds for the follow-up experimental triplicate dose-response confirmation. They identified 17 compounds to possess effectiveness including *Remdesivir*, of which 10 are novel *in vitro* identifications.

Heiser *et al.* [38] used a morphological analysis pipeline one RxRx19a dataset, a fluorescence microscopy dataset of HRCE samples, to identify potential treatments for COVID-19. The complete workflow is not outlined in their paper, since they are using proprietary pipelines. They used a proprietary deep convolutional neural network to calculate dense embeddings for the sample images. These embeddings ar then used for calculating on-disease projection and off-disease rejection scores per dose for each treatment. These scores were

then aggregated and normalized using a proprietary algorithm to compute the hit-scores. They reported that *Remdesivir*, *GS-441524*, *Aloxistatin*, *Silmitasertib*, and *Almitrine* showed moderate to strong effectiveness in their model, whereas, neither *chloroquine* nor *hydroxychloroquine* demonstrated any significant effectiveness in their cultured cells. Similarly, Cuccarese *et al.* [22] used an *in vitro* deep-learning-driven analysis of cellular morphology on HUVEC cells with COVID-19-associated cytokine storm to identify potential candidates for drug repurposing.

2.3 Multiple Instance Learning

Multiple instance learning (MIL), as a form of weakly supervised learning, has been under the spotlight of research communities recently due to its ability to leverage weak supervision for tasks that are conventionally considered heavily dependent on laborious human annotations [11]. MIL was originally introduced for drug activity prediction [26] and recently it has been applied in many different domains such as computer vision [94], medical imaging and diagnosis [101, 10, 107], and in biology for applications such as mechanism of action classification using microscopy images [50], identifying antigen binding peptides [108], predicting specific functional binding sites in microRNA targets [7] to name a few. Contrary to conventional supervised learning methods in which every instance in the training data is associated with a label, in MIL the learner is provided with a training set of labeled bags, where each bag is a set of instances. The learner is tasked to predict a label for an unseen bag given its instances. MIL algorithms can be categorized into three classes regarding how and on what level they extract their required discriminative features [4]:

- **Instance-space** algorithms are based on the assumption that the information needed for the bag classification lies within instances and are local. Hence, an instance-level classifier is trained in the first place, and then an aggregation method is used to integrate the instance-level predictions into a single bag-level label. There are multiple assumptions about the relation between the instance-level and bag-level labels [30]. The most commonly used assumption is the standard multiple instance assumption, which has been incorporated in methods such as EM-DD [112], miSVM [5], MIBoosting [104], and VF [57]. This assumption states that in a binary bag classification configuration, each bag labeled as positive class contains at least one instance representing the positive class, whereas the negative class bags do not contain any instance of positive class [60].

- In **Bag-space** algorithms, the learner extracts the discriminative information from the whole bag as a composite entity of instances based on the assumption that the discriminative features exist at the bag level, and predicts a label for the bag using distance or similarity using classifiers such as K-NN or SVM in MInd and mi-Graph

[15, 116]. Contrary to instance-space paradigm that uses local instance-level information, this class of algorithms extracts global information by taking the whole bag into account.

- In the **Embedded-space** class, the learner first maps all instances in the bag into a vector in the embedding space, and then a classifier is trained on this embedding space. Similar to the bag-space algorithms, the embedded-pace uses the global bag-level information to learn the discriminative features for classification, however, the difference between these two approaches resides in the feature extraction. While in the bag-space methods the information is extracted implicitly via the distance or similarity function, in the embedded-space algorithms this is done explicitly by the embedding function that maps the bags into a vector space. This approach converts the multiple instance classification problem into a conventional supervised classification one.

2.4 Deep Multiple Instance Learning

Considering the prevalence of deep neural networks and MIL ability to formulate and tackle problems with weak supervision, the integration of both approaches was unsurprising. Deep multiple instance learning uses a deep neural network model as a learner in the MIL framework, and it has shown performance competitive with state-of-the-art in recent studies. Hou *et al.*, [40], reported an improved performance of a MIL based instance classifier with extracted patches from whole slide tissue images for predicting the subtypes of glioma and non-small-cell lung carcinoma cases. Campanella *et al.* [10] proposed an embedded-space pipeline for classifying and finding potentially cancerous candidate regions on whole slide tissue images by training an patch classifier coupled with a recurrent neural network to embed and aggregate patch information for whole slide classification. They used an instance classifier based on the standard MIL assumption to classify the extracted patches of the whole slide tissue images, and then used the penultimate layer of the trained model as an embedding for each instance. A recurrent neural network takes a bag in the form of a sequence of its instances' embeddings and aggregates them for the final bag-level label prediction. Similarly, Ilse *et al.* [43] proposed a fully trainable deep MIL model combined with an attention mechanism [103] to associate a learnable weight to each instance or its embedding and evaluated it on two datasets of histology images for breast and colon cancer. Applying the attention mechanism in the MIL context has been very limited in the literature and due to the attention mechanism's nature, it can only be applied to the embedded-space methods [72]. In this context, the attention mechanism is integrated into the pooling function that aggregates the instances or their lower dimension embeddings into a fixed-size tensor before predicting the bag label. In contrast, when using instance-space methods, attention mechanism is not applicable since the instance aggregation is disjoint from instance classifier training.

Kraus *et al.* [50] introduced a new architecture for classification and segmentation of microscopy images using Noisy-AND pooling layer for aggregation of instances and showed improved performance for mechanism of action classification on mammalian and yeast datasets. Yan *et al.* [105] proposed a dynamic permutation-invariant pooling function for deep MIL architectures that estimates each instance contribution to its bag label iteratively and reported improved performance on 4 different classification tasks on various datasets. Recently, Zhao *et al.* [114] proposed a embedded-space deep MIL based on a graph neural network for predicting lymph node metastasis in colorectal cancer on histopathological images. Their model consists of variational autoencoders and generative adversarial networks to extract the discriminative features and embed the bags into a latent space. A graph convolutional network is then applied on top of the extracted embeddings to predict the lymph node metastasis.

2.5 Contribution

Virus-specific markers are not always included in high-throughput imaging studies that investigate the effects of viral infection and drug treatments on a host cell morphology; in these instances, the extent of viral infection is unknown. This scenario may also occur during the very early stages of an outbreak or pandemic when limited virus specific tools are available, such as when COVID-19 first emerged. When no pixel-level annotations are available to specifically identify infected regions of the imaged cell population, MIL may be used as an alternative strategy to localize viral infection and the corresponding effects on host cell morphology. The imaged cell populations can be broken down into smaller instances, single cells or patches, and the sample cell population label is utilized in the MIL formulation. An instance-space approach would very well suit this problem for viral infections such as SARS-CoV-2, that induce local morphological changes to cells. Based on the standard assumption, we assume that if a sample cell population is infected, then there should be at least one instance that contains infected cells. And a non-infected sample cell population does not have any infected cells.

In this work, we present DEEMD: a computational pipeline that estimates treatment efficacy using morphological features of cells extracted using a deep learning model. DEEMD is scalable to process and screen multiple cell lines and thousands of treatments in parallel and can accelerate the identification of clinically evaluated and FDA-approved compounds with antiviral activity. We applied DEEMD on a public microscopy image dataset of SARS-CoV-2 infected cells, and it identified treatment compounds that have been suggested in the literature to be effective against COVID-19, further supporting the validity of the pipeline. The specific design of the deep learning model allows us to localize infected cells via a weakly supervised training procedure without expensive annotations. Localizing the infected

cells allows us to better explain the model predictions, as well as enhances our biological understanding of the morphological changes in cells induced by SARS-CoV-2.

To the best of our knowledge, this is the first work to approach the drug repurposing problem via leveraging weak supervision through multiple instance learning. Many morphological analysis pipelines utilize a viral-specific stain to localize infected regions of interest in sample images for studying morphological changes induced by the viral. However, these viral-specific stains are expensive and in case of an emerging virus, unavailable to be utilized. Another approach taken by other morphological profiling pipelines, is to rely on computer vision models to extract viral-related morphological features through massive datasets without providing any annotations on infected regions of interest. DEEMD is an attempt to integrate both these approaches through identifying infected regions in sample images without requiring any viral-specific stains or annotations.

DEEMD could be used in the very early stages of an outbreak to quantify the effectiveness of any known and approved compound and identify lead small molecules or compound candidates that show therapeutic effectiveness against viral infection. The identified candidates can then be used in more targeted testing once other molecular tools become available and that the findings can be used in combination with other methods, such as fluorescent-tagged virus, nucleic acid-based methods (PCR), or animal models.

Chapter 3

Methodology

We first describe a high-level overview of the DEEMD before delving into the details. DEEMD consists of two components: First component is a deep learning model which is trained on microscopy images of untreated samples to extract discriminative morphological features in MIL framework to distinguish between uninfected and SARS-CoV-2-infected cell population images, further described in Section 3.1. Following infection of cells treated with a drug, the trained deep learning model is used to extract a morphological profile for each sample. These profiles are then aggregated and used in a statistical model, the second component, to estimate the treatment efficacy per concentration, Section 3.4. Finally, all dose-dependent efficacy scores are aggregated for each treatment to form the final identified set of potentially effective treatments. In this context, we define treatment effectiveness as how close a treated sample's morphological profile is to uninfected cell profiles. We discuss each component detail in the following sections.

3.1 Classification with deep MIL

In this work, we follow closely the training procedure described in [10, 105] for training an instance classifier based on a relaxed version of the standard multiple instance assumption. The training set $\mathcal{D}$ consists of n data points $(\mathbf{X}, \mathbf{Y})$ in form of (x_i, y_i) pairs where x_i is a fluorescence microscopy image along with its associated sample-level label y_i with two classes: active SARS-CoV-2 (c_P) and non-infected (c_N). Each sample image x_i can be split into N patches each referred to as $x_i^{p_j}$. We assume that each $x_i^{p_j}$ is associated with a patch-level label $y_i^{p_j}$ which is unknown and not included in the dataset $\mathcal{D}$. Let $\mathcal{M}(\theta)$ be a deep neural network model responsible for predicting a patch-level label $\mu_i^{p_j}$ for every patch $x_i^{p_j}$ of each sample image x_i. In the MIL terminology, model $\mathcal{M}(\theta)$ is an instance classifier since it classifies each instance into active SARS-CoV-2 and non-infected classes. We define $\mu_i^{p_j}$ to be the model $\mathcal{M}(\theta)$ estimation of the unknown true patch-level label $y_i^{p_j}$ given the dataset

$\mathcal{D}$. This can be expressed mathematically as:

$$\mu_i^{p_j} \triangleq \mathbb{P}\Big[y_i^{p_j} = c_P \Big| \mathcal{D}; \mathcal{M}(\theta)\Big], \tag{3.1}$$

where $\mathbb{P}[A]$ shows the probability of the event A.

Each training iteration starts with an exhaustive inference over all patches in the training set and all corresponding $\mu_i^{p_j}$ are estimated. Then for each sample image x_i, all of its patch-level estimated labels $\mu_i^{p_j}$, denoted as M_i, are sorted and the set of k patches with the highest probability, denoted as $\mathcal{K}_i^k$, are selected for model training.

$$t_r \triangleq \Big\{ j \Big| \mu_i^{p_j} = m(M_i, r) \Big\}, \quad \mathcal{K}_i^k \triangleq \bigcup_{r=1}^{k} t_r, \tag{3.2}$$

where we define $m(S, r)$ to be the r-th greatest element in the set S.

The sample-level label y_i is assigned to each set of patches in $\mathcal{K}_i^k$, and the model $\mathcal{M}(\theta)$ is trained to minimize the binary cross entropy loss, $\mathcal{L}\big(\cdot\,|\mathcal{D}; \mathcal{M}(\theta)\big)$, between y_i and $\mu_i^{p_j}$ for all patches in the set $\mathcal{K}_i^k$. More specifically, we can break this process down into multiple steps:

1. Given a patch p_j and the label y_i associated with its corresponding sample image x_i, we can calculate binary cross entropy loss as follows:

$$\mathcal{L}(x_i^{p_j} | \mathcal{D}; \mathcal{M}(\theta)) = -y_i \log(\mu_i^{p_j}) - (1 - y_i) \log(1 - \mu_i^{p_j}). \tag{3.3}$$

2. We define the loss value for the sample image x_i to be only inclusive of informative patches, e.g. patches that we have confidence in their predicted infection probability. In other words, we only include patches which are in the set $\mathcal{K}_i^k$ and take an overage over them. Thus the value of the loss function for each sample image is defined as follows:

$$\mathcal{L}(x_i | \mathcal{D}; \mathcal{M}(\theta), w_P, w_N) = -\frac{1}{k} \sum_{j \in \mathcal{K}_i^k} \Big[w_P y_i \log(\mu_i^{p_j}) + w_N (1 - y_i) \log(1 - \mu_i^{p_j}) \Big], \tag{3.4}$$

where w_P and w_N are incorporated in the loss function $\mathcal{L}\big(\cdot\,|\mathcal{D}; \mathcal{M}(\theta)\big)$ to account for class imbalance [46].

$$\mathbf{Y}_P = \Big\{ y_i \Big| y_i = c_P, i \in \{1, \cdots, n\} \Big\}, \quad \mathbf{Y}_N = \Big\{ y_i \Big| y_i = c_N, i \in \{1, \cdots, n\} \Big\}, \tag{3.5}$$

$$[w_P, w_N] = \frac{|\mathbf{Y}|}{|\{c_P, c_N\}|} \times \Big[\frac{1}{|\mathbf{Y}_P|}, \frac{1}{|\mathbf{Y}_N|} \Big], \tag{3.6}$$

where $|A|$ represents the cardinality of A.

3. Finally the loss function $\mathcal{L}(\cdot|\mathcal{D};\mathcal{M}(\theta))$ for the whole dataset in each epoch is calculated as follows:

$$\mathcal{L}(\mathbf{X}|\mathcal{D};\mathcal{M}(\theta),w_P,w_N) = -\frac{1}{n}\sum_{i=1}^{n}\mathcal{L}(x_i|\mathcal{D};\mathcal{M}(\theta),w_P,w_N). \tag{3.7}$$

The reader is encouraged to refer to Section 5.6 for a detailed discussion on how hyper-parameter k affects the training dynamics for a MIL learner.

3.2 Inference with deep MIL

Since the classification labels are only available for the sample images, aggregation of the patch-level infection probabilities is required for performance evaluation and down-stream analysis. For inference, each sample image is split into patches and a forward pass through the model. This provides us with the estimated infection probabilities, $M_i = \Big\{\mathbb{P}[y_i^{p_j} = c_P|\mathcal{D};\mathcal{M}(\theta)]\Big\}_{p_j\in x_i}$, which are aggregated to form the sample-level label $\hat{y}_i$:

$$\hat{y}_i = \mathbb{I}\Big(\min_{p_j\in\mathcal{K}_i^k}\{\mu_i^{p_j}\} \geq \eta\Big) = \mathbb{I}\Big(m(M_i,k) \geq \eta\Big), \tag{3.8}$$

where η is the cut-off threshold, selected based on the validation set and $\mathbb{I}(\cdot)$ is an indicator function.

For the MIL model $\mathcal{M}(\theta)$, hyper-parameter k is incorporated into the standard multiple instance assumption so that the model can generalize to a more relaxed constraint: a positive (SARS-CoV-2 infected) sample image should contain at least k positive patches to be considered positive and a sample image would be considered negative if less than k patches are predicted to be infected. This generalized assumption is mathematically expressed in Eq. 3.8:

3.3 MIL infection localization

Localization of the infected region is achieved using the estimated $\mu_i^{p_j}$. To form the infection map A_i all patches need to be aggregated. Based on the set of all patches that overlap at the pixel $x_i^{(l,m)}$, denoted as set $\mathcal{O}_{x_i}(l,m)$, the value of infection map $A_i^{(l,m)}$ is calculated by the weighted average of its overlapping patches infection probabilities, $\{\mu_i^{p_j}\}_{p_j\in\mathcal{O}_{x_i}(l,m)}$. We opt to set sample weights based on the sample's values, similar to the quadratic mean, but using $(\mu_i^{p_j})^\alpha$, where $\alpha < 1$, instead of $\mu_i^{p_j}$. The infection map A_i is calculated as:

$$A_i^{(l,m)} = \frac{\sum_{p_j\in\mathcal{O}_{x_i}(l,m)}(\mu_i^{p_j})^{1+\alpha}}{\sum_{p_j\in\mathcal{O}_{x_i}(l,m)}(\mu_i^{p_j})^{\alpha}}.$$

By using this averaging method, with $\alpha < 1$ when all $\mu_i^{p_j}$'s are smaller than 1, the average is more sensitive to higher values, and smaller values have less diminishing power on the result compared to simple averaging. This property is well-suited for infection localization since patches with high infection probabilities do not get diminished by the adjacent overlapping low probability ones. Finally, prior to rendering the infection map, a low-pass Gaussian filter is applied to make the infection map smoother. For generating the infection map A_i based on the MIL model, $\mathcal{M}(\theta)$ we set α to 0.2.

3.4 Treatment Efficacy Estimation

DEEMD estimates the treatment efficacy based on its trained MIL model that compares the morphological profile of drug-treated cells with those of untreated infected and uninfected cells. We assume that an effective treatment would prevent drastic infection-induced morphology changes in the cells, hence, the treated cells morphological profile would be similar to that of uninfected ones. A statistical model takes in the infection probabilities predicted by the MIL model and estimates the probability that a treatment is effective based on its morphological similarities to uninfected cells.

For each sample image x_i, the infection probability z_i is calculated as:

$$z_i = \text{Median}\{\mu_i^{p_j} | p_j \in \mathcal{K}_i^k\}. \tag{3.9}$$

We define the set $\mathcal{T}_{t_i}^{c_j}$ such that it consists of infection probabilities for every sample image of a given treatment t_i at concentration c_j in the treated test set. We have to aggregate the infection probabilities, z_i, over the set $\mathcal{T}_{t_i}^{c_j}$ to minimize the inevitable variations that have occurred during the sample preparation, treatment administration and image acquisition. We define $e_{t_i}^{c_j}$ to be the dose-dependent estimated efficacy score of the treatment t_i at concentration c_j based on any model $\Omega(\omega)$ capable of assigning an infection probability to a sample image based its morphology. Since we do not have any assumptions on the distribution of infection probabilities z_i in the set $\mathcal{T}_{t_i}^{c_j}$, we opt to use non-parametric statistics related to the median of it, $\beta_{t_i}^{c_j}$. We observed that the distribution infection probabilites z_i in the set $\mathcal{T}_{t_i}^{c_j}$ for any t_i is heavily skewed and asymmetric, thus using Wilcoxon test is misleading since it assumes a symmetric distribution for the data[98]. However, the sign test is still valid[27]. Given a confidence level a, we can calculate the exact confidence interval for the estimated median, denoted as $\text{CI}(\beta_{t_i}^{c_j} | a, \Omega(\omega))$. To be more conservative about the false positive rate, instead of using the point estimate $\beta_{t_i}^{c_j}$, we use the least upper bound, supremum, of $\text{CI}(\beta_{t_i}^{c_j} | a, \Omega(\omega))$. The mathematical expression for the dose-dependent estimated efficacy score of the treatment t_i at concentration c_j is:

$$e_{t_i}^{c_j} \triangleq T(\mathcal{T}_{t_i}^{c_j} | \Omega(\omega)) = 1 - \sup\{\text{CI}(\beta_{t_i}^{c_j} | a, \Omega(\omega))\}, \tag{3.10}$$

where $T(\cdot)$ is a descriptive statistic. The dose-dependent efficacy score $e_{t_i}^{c_j}$ reflects the model $\Omega(\omega)$ belief on the morphological similarity between the sample images in the set $\mathcal{T}_{t_i,c_i}$ and the uninfected ones in the training set. Close to 1 values for $e_{t_i}^{c_j}$ imply that the treatment t_i at concentration c_j is effective against SARS-CoV-2. In this context, we assume that the morphological profile of cells treated with an effective treatment is more similar to the uninfected cells rather than infected ones.

To summarise all estimated dose-dependant efficacy scores of a given treatment t_i for raking all treatment compounds and identifying the effective ones, the estimated efficacy score e_{t_i} is calculated:

$$\tilde{e}_{t_i}^{c_j} = \begin{cases} \{e_{t_i}^{c_j} | e_{t_i}^{c_j} \geq \zeta\}_{c_j \in \mathcal{C}_{t_i}} & \exists c_j : e_{t_i}^{c_j} \geq \zeta \\ \{e_{t_i}^{c_j}\}_{c_j \in \mathcal{C}_{t_i}} & \text{otherwise,} \end{cases} \tag{3.11}$$

$$e_{t_i} = \text{Median}\{\tilde{e}_{t_i}^{c_j}\} \tag{3.12}$$

where ζ is a cut-off threshold and $\mathcal{C}_{t_i}$ refers to all available concentrations of treatment t_i in the treated test set. The treatment compounds that DEEMD identifies as effective against SARS-CoV-2 are those for which at least one dose-dependent efficacy score is higher than ζ or equivalently: $\mathcal{E}_{\mathcal{M}(\theta)} = \{t_i | e_{t_i} \geq \zeta\}$.

Chapter 4

Experimental Design and Implementation Details

Here we describe the dataset that we used in this work along with the implementation and design details needed for reproducing the presented results, and adapting DEEMD to other domains and problems.

4.1 Dataset

We are using the publicly available RxRx19a dataset, which is the first morphological imaging dataset of cells infected with SARS-CoV-2 [38]. The dataset consists of human renal cortical epithelial (HRCE) and African green monkey kidney epithelial (VERO) cell lines subjected to three different conditions: 1) mock uninfected control, 2) infection with ultraviolet (UV)-light inactivated SARS-CoV-2 (irradiated for 20 minutes), or 3) infection with SARS-CoV-2 at multiplicity of infection of 0.4, as illustrated in Fig 4.1. All samples were incubated for 96 hours. A library of 1672 small molecules and FDA-approved treatments were applied to a subset of active SARS-CoV-2 infected samples in 6+ half log concentrations with six replicates per dose for each compound.

All cells were stained with five fluorescent dyes detecting various subcellular structures, each imaged in a separate channel on a fluorescent confocal high-content imaging microscope. The dyes include Hoechst (nucleus), Syto14 (nucleoli and cytoplasmic RNA), phalloidin (actin cytoskeleton), Concanavalin A (ConA; endoplasmic reticulum) and Wheat Germ Agglutinin (WGA; Golgi and plasma membrane). In Table 4.1, a representative set of sample images from each condition. The dataset consists of more than 300K 5-channel labeled images of size 1024×1024 and is publicly available through the *Recursion* website: `www.rxrx.ai/rxrx19a`.

For this work, we only used the HRCE cell line and merged the mock control and the UV inactivated control into a single uninfected class, as no cytopathic effects were observed in either condition [38]. We also verified this by using each sample image's corresponding

deep learning embeddings included in the dataset. We decided to merge mock control and UV inactivated into a single class as they were indistinguishable by a trained classifier (data not shown).

4.2 Cell nucleus count in the sample images

RxRx19a dataset includes Hoechst stain which targets nucleus. We used this stain to identify and count the cells in each sample image for both explanatory analysis and preprocessing the data. We first stitch images of 4 adjacent sites of each well on each plate to reconstruct the image of whole sample cell population in each well. This step is necessary to avoid nuclei multiple counting. We localize and count the nuclei in each sample image using a segmenting pipeline based on Otsu thresholding. We used this pipeline to exclude any sample images which did not contain any detectable cells from the dataset. The distribuion of the nuclei count in the sample images for infected and non-infected samples are shown in Fig 4.2. The cell nucleus count is significantly lower in the infected samples (using two-sided Mann-Whitney test with $p = 0.0$).

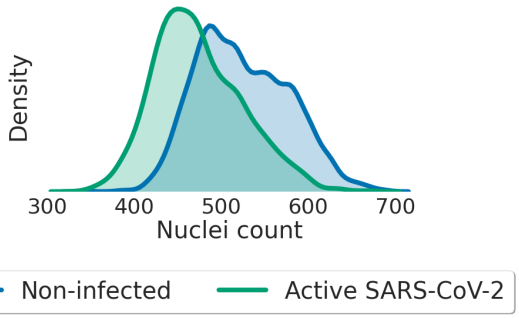


Figure 4.2: Distribution of the nuclei count in the stitched well sample images. The non-infected class consists of samples from both Mock and UV Inactivated SARS-CoV-2 classes.

4.3 Baseline Classification Models

To compare the MIL model $\mathcal{M}(\theta)$ performance, two other deep learning model are trained: A model that is trained with the conventional training procedure for training convolutional neural networks by minimizing a cross-entropy-based loss function, similar to Eq. 3.4, with full resolution sample images as input. Denoted as $\mathcal{W}(\phi)$, this model will be referred to as whole-image based model $\mathcal{W}(\phi)$ throughout this thesis. Unlike most deep neural network classifiers, the input sample image is not down-sampled, and the full resolution version was used to make the comparison fair. The whole-image based model $\mathcal{W}(\phi)$ takes in the sample image x_i and calculates infection probability for that sample population.

Instead of limiting the model to only use k instances with the highest probabilities within each bag for training, we can have a reasonable alternative approach of using all instances in the bag and assigning them with the bag-level label. There are studies in histopathology such as [19] that used this alternative approach and their proposed model outperformed other models at their respective tasks. We refer to this model as the patch-based model $\mathcal{V}(\psi)$ in the text. Similar to MIL instance classifier, $\mathcal{V}(\psi)$ calculates a patch infection probability

for all patches in the input sample image x_i. Intuitively, in this approach the model would potentially be able to eliminate the effect of noisy labels and estimate the true distribution of the instance labels given large enough training data. This approach is an extreme case of the MIL training procedure in which $k = N$ is where N is the total number of instances within a bag.

For evaluating the patch-based model $\mathcal{V}(\psi)$, an averaging method is used for aggregating the patch-level estimations:

$$\hat{y}_{i,\mathcal{V}(\psi)} = \mathbb{I}\Big(\frac{\sum_{j\in\{1,\dots,N\}} \delta_i^{p_j}}{N} \geq \eta\Big) \tag{4.1}$$

Where η is the cut-off threshold, selected based on the validation set, and $\mathbb{I}(\cdot)$ is an indicator function.

4.4 Implementation Details

We used a 5 cross-fold validation approach for training and testing each model. The HRCE sample images in RxRx19a dataset are split into 4 non-overlapping sets: **1)** training, **2)** validation, **3)** untreated, and **4)** treated test sets consisting of 20K, 5K, 6K, and 247K sample images respectively. The training set is used for updating the weights and training the neural networks using backpropagation, whereas the validation set is used to tune hyper-parameters. The performance of the tuned model is measured on the held-out untreated test set, while the treated test set is solely used for estimating the efficacy of the treatments. All sets are balanced in terms of class labels except for the treated test set which only consists of active SARS-CoV-2 infected sample images.

All models are based on ResNet34 architecture, pretrained on ImageNet with a modified input layer to accommodate 5-channel inputs [37, 25]. Each model was trained for 150 epochs, monitored for early stopping, with Adam optimizer at a learning rate $\eta = 10^{-4}$, $\beta_1 = 0.9$ and $\beta_2 = 0.999$, [47], with a batch size of 128. All hyper-parameters were tuned with respect to the model's performance on the validation set. Sample images are channel normalized using the empirical means and standard deviations calculated on the training set. We used $\eta = 0.5$ for evaluating the performance of the models. For the treatment efficacy estimation, ζ was set to 0.5 and a confidence level of 0.95% was used in the sign test. We used *PyTorch* framework [73] for model training and evaluation and treatment efficacy estimation was coded using *R* [79].

For both the MIL model $\mathcal{M}(\theta)$ and the patch-based model $\mathcal{V}(\psi)$, we used a uniform grid of 256×256 patches with 50% overlap, resulting in 49 patches per sample image. Since we wanted DEEMD to be applicable to other microscopy imaging datasets with minor modifications, we choose a uniform grid of patches over a cell-based segmentation map. Although a cell-based segmentation map could potentially provide a single cell resolution

infection map, it is heavily dependent on the stains used in the imaging procedure, which in turn reduces the pipeline generalization. Based on the validation set, we set the hyper-parameter k for the MIL model $\mathcal{M}(\theta)$ to 2. The whole-image based model $\mathcal{W}(\phi)$ was trained on full-resolution (1024×1024) sample images without any input downsampling.

We trained the models on a single Nvidia *TITAN V* GPU. Notice that the effective training set for the MIL model $\mathcal{M}(\theta)$ on each training iteration is k/N of the whole training samples, thus its training is N/k times faster compared to the patch-based model $\mathcal{V}(\psi)$. The patch-based model $\mathcal{V}(\psi)$ even has longer training time than the whole-image based model $\mathcal{W}(\phi)$ since the patches are overlapping. Thus, the MIL model $\mathcal{M}(\theta)$ is computationally more efficient compared to the other two.

4.5 Selecting optimal value for hyper-parameter k

Choosing an optimal value for hyper-parameter k is crucial to the performance of MIL models since it effectively controls the amount of noise in the training labels. The hyper-parameter k is the number of top high probability instances in each bag used for loss calculation in each training iteration, 3.4. Tuning this hyper-parameter can be either based on problem-specific domain knowledge or a performance metric on the validation set. For the problem of estimating the treatment efficacy against SARS-CoV-2, the choice of optimal k was found based on incorporation of domain knowledge and common model performance metrics.

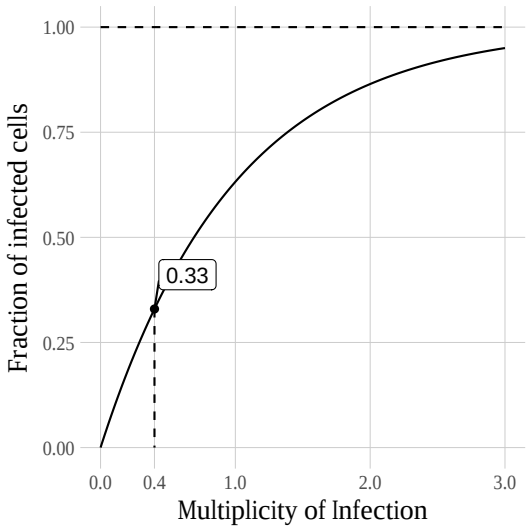


Figure 4.3: Cell infection probability as a function of MOI in a simple modeling.

In a simple yet reasonable modeling, let's assume that the virions enter the cells independent of each other and X is a random variable that counts the number of virions entering a cell. Since we are dealing with a counting process, it is reasonable to assume that random variable X comes from a Poisson distribution with parameter m, which is actually the multiplicity of infection or MOI $X \sim \text{Poi(m)}$. The probability of a cell being infected is calculated as follows:

$$\mathbb{P}[\text{A cell is infected}] = 1 - \mathbb{P}[X = 0] = 1 - e^{-\text{m}} \tag{4.2}$$

The probability of a cell being infected is shown in Fig 4.3 as a function of MOI. Based on the data provided in [38], in the preperation process of the the RxRx19a dataset, the samples were infected with a MOI of 0.4. Based on our modeling of the infection process, this value would indicate that around 33% of the cells in each sample are infected on average.

We incorporate this domain knowledge about the expected number of infected cells within each sample image in selecting the optimal value for hyper parameter k. Since we are using a uniform grid of overlapping patches for the sample images, the number of the selected patches in the set $\mathcal{K}_i^k$can not be directly translated into the fraction of the infected cells within each sample image. The area of the infected regions based on the infection maps can be a good approximation for the faction of the infected imaged cells. We trained multiple MIL models with different values for the hyperparemeter $k \in \{1, 2, 3, 5, 10, 15, 25, 49\}$. Next, we calculated the fraction of pixels that were higher than the cutoff threshold η in the generated infection map for each sample images in the validation set for all of the MIL models, as shown in Fig 4.5. We can see that when k equals to 1 or 2 or 3, the average of the distribution is close to the theoretical calculated value using the MOI. Hence, these 3 options are candidates for the optimal value for k. Finally, to select the optimal value between these 3 candidates, we turn to the average precision metric. As shown in Fig 4.4, we can see that the models have a very close AUC, but when K is 2 the performance is slightly better and thus we set k to be 2.

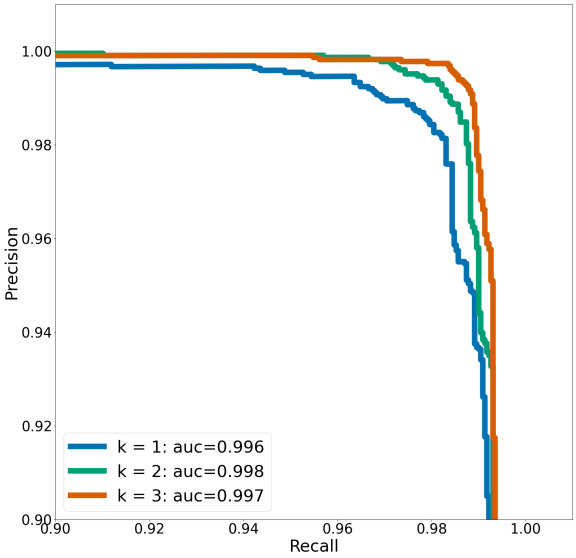


Figure 4.4: Precision-Recall curve for candidate optimal values of k

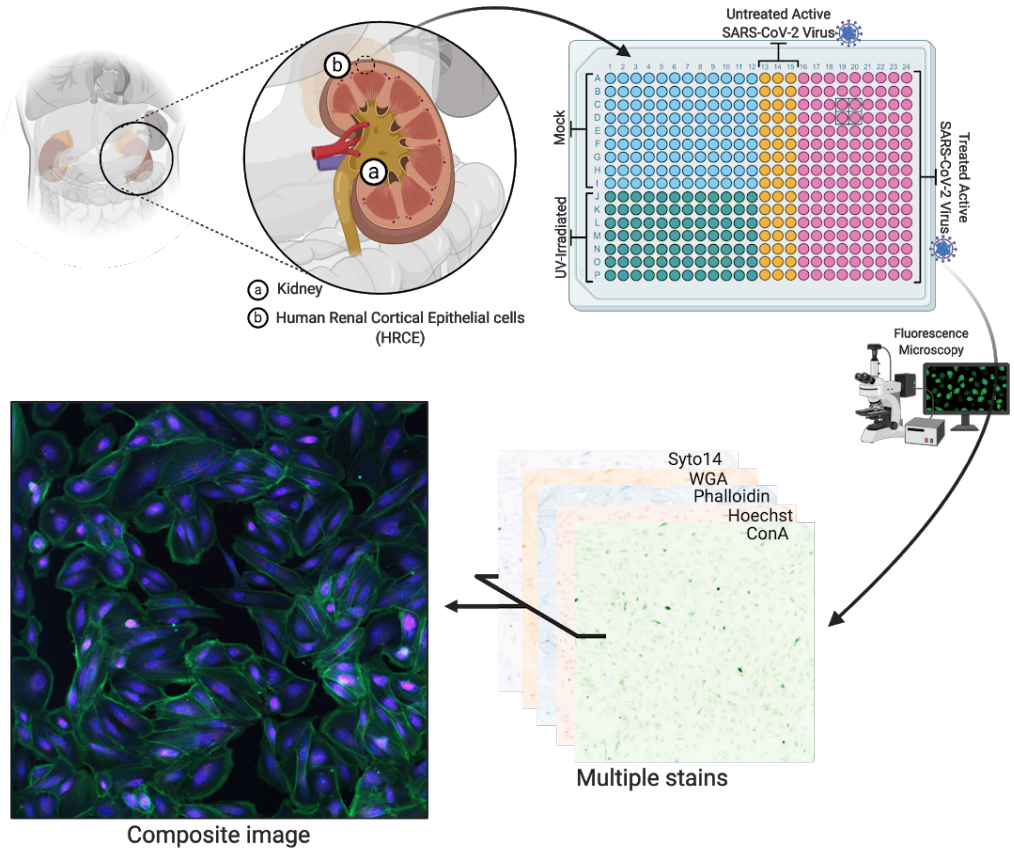


Figure 4.1: RxRx19a dataset includes monolayers of normal HRCE cells in three viral conditions: Mock, UV Inactivated SARS-CoV-2, Active SARS-CoV-2. After incubation for 96 hours, the samples were fixed and stained with 5 dyes and imaged on a fluorescent confocal microscope.

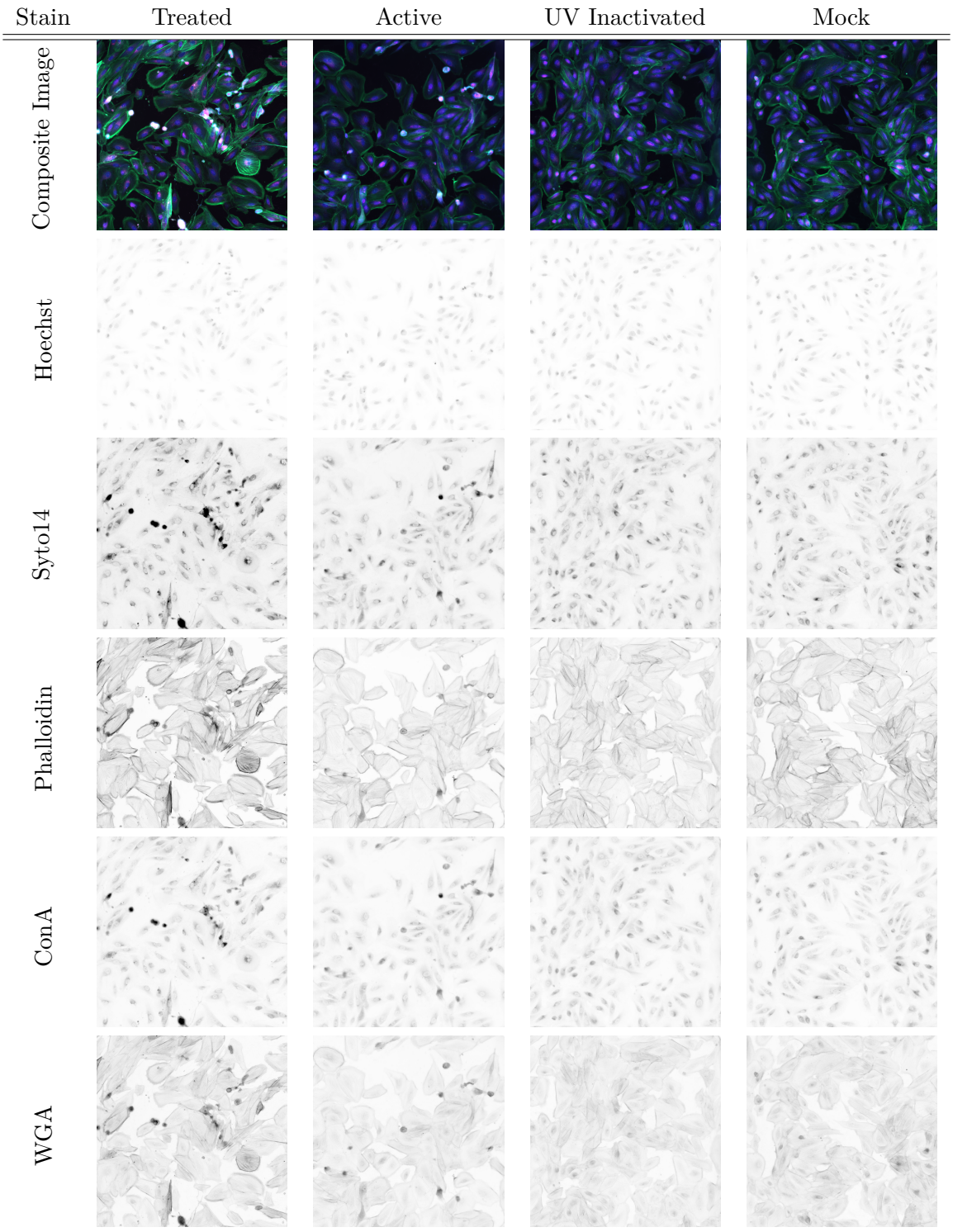


Table 4.1: Representative sample images from each condition in the RxRx19a dataset: Mock, UV Inactivated SARS-CoV-2, Active SARS-CoV-2, and Treated Active SARS-CoV-2. Cells were fixed and stained with fluorescent dyes each detecting a specific subcellular structures and imaged on a fluorescent confocal high-contentimaging microscope. For better visualization, each dye image is color-inverted. The composite image is constructed by overlaying each stain in a color channel: Hoechst (violet), Syto14 (red), Phalloidin (green), ConA (blue), and WGA (cyan).

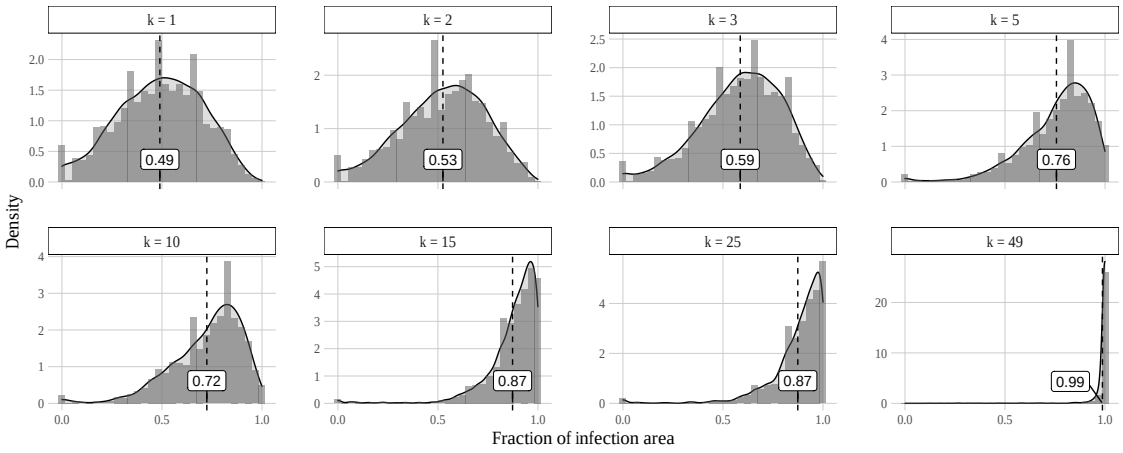


Figure 4.5: The distributions of the fraction of infected pixels in each sample image in the validation set for multiple values of k, along with the mean of the distribution.

Chapter 5

Results and Discussion

5.1 The MIL model can accurately predict SARS-CoV-2 infection

We evaluate the performance of the MIL model $\mathcal{M}(\theta)$, along with two baseline models, on the untreated test set in terms of the area under the curve for precision-recall curve, or average precision, as shown in Fig 5.1. The precision-recall curve effectively represents the trade-off between precision and recall for all possible cut-off values based on the model predictions. All three models are capable of accurately classifying the sample images into non-infected and SARS-CoV-2 infected classes, with an average precision of $\approx$ 0.99. The performance of the models is independent of the learner architecture; similar results were observed with models based on VGG16 architecture [85].

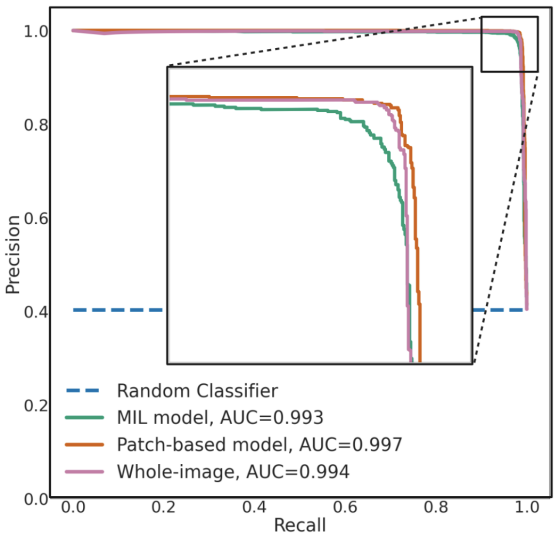


Figure 5.1: Comparison of precision-recall curves. All three models are capable of accurately classifying the sample images in the untreated test set. A random classifier is included as a reference.

High average precision implies that these models have learned morphological features that can be generalized to the untreated test set for accurate classification, however, the learned feature spaces have drastically different characteristics due to their input and training procedures. The MIL model $\mathcal{M}(\theta)$ has learned to extract highly discriminative features from micro-populations or single cells whereas the whole-image based model $\mathcal{W}(\phi)$ focuses on the macro-population

of cells and learns morphological features of the population, which are not as detailed as the MIL model $\mathcal{M}(\theta)$'s. The patch-based model $\mathcal{V}(\psi)$ lies in between those two models in the spectrum; on the one hand it has been trained using patches with micro-population features, but on the other hand, the labels it was provided for training were noisy. Thus, this model is not focused on the details as the MIL model $\mathcal{M}(\theta)$. These shifts in the learned feature spaces are further transferred to the downstream analysis of treatment efficacy, as discussed in Section 5.3 and Section 5.5 .

5.2 Visualization of the MIL model $\mathcal{M}(\theta)$ morphological feature space

To better understand the MIL model $\mathcal{M}(\theta)$ learned morphological feature space, we visualized the penultimate layer in its ResNet34 architecture. We trained UMAP dimensionality reduction [61] on the training set and use it to project the patched sample images from the untreated test set to get the 2-dimension projections. The calculated projections are demonstrated in Fig 5.2. The original patches are also overlaid on top of the points to further demonstrate the morphological differences between the two classes and the transition between the two. We can see that there is a rather smooth transition from the noninfected samples to the infected ones. The RxRx19a dataset does not include a fluorescent marker specific for SARS-CoV-2 that can verify which individual cells in a sample image are infected, thus the extent of infection is not known and cannot be verified. However, the fraction of patches predicted as being infected by the MIL model $\mathcal{M}(\theta)$ is consistent with a SARS-CoV-2 MOI of less than one used in the preparation of the RxRx19a dataset. We further explain the morphological features that we hypothesize the MIL model $\mathcal{M}(\theta)$ is sensitive to in Section 5.8.

5.3 DEEMD dose-dependent efficacy scores are well-structured

We assume that an efficacious treatment is able to effectively stop viral infection and prevent major infection-induced morphological changes in the cell population. Using morphological analysis we can estimate treatment efficacy by profiling the treated and infected cell morphology and quantifying its similarity to non-infected and infected morphology. We applied DEEMD to the infected and treated cell images from the RxRx19a dataset which resulted in a ranked list of potential efficacious treatments against SARS-CoV-2, $\mathcal{E}_{\mathcal{M}(\theta)}$ shown in Fig 5.3.

The dose-dependent efficacy scores for top ranked effective compounds identified by DEEMD are shown in Fig 5.4-(a). As is expected for effective antiviral compounds, the dose-dependent efficacy score of the identified treatments increases with increasing concentration, similar to the fitted logistic curve. We expect to observe higher effectiveness with increasing

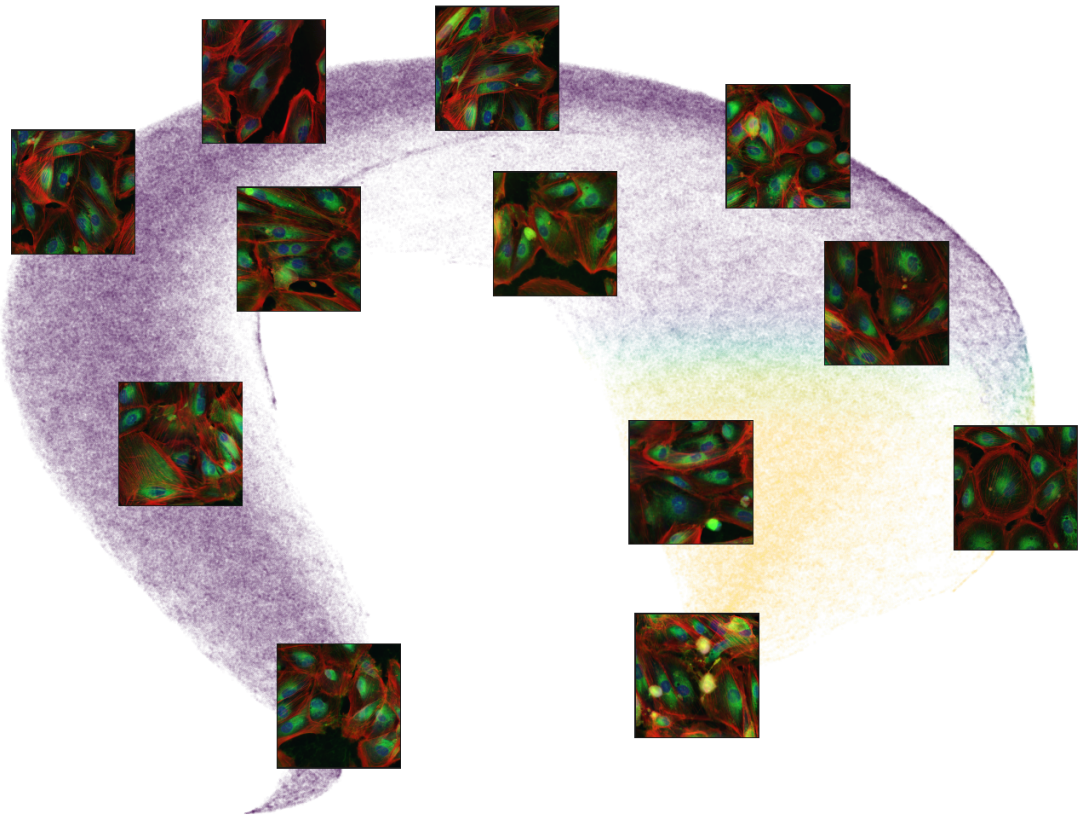


Figure 5.2: Visualization of the MIL model $\mathcal{M}(\theta)$ learned morphological space. Each data point is a patch from the sample images in the untreated test set. Some of the original patches are overlaid for better visualization of the transition. The projected points are color coded: purple indicates low probability of being active SARS-CoV-2 and, by transitioning into yellow, the infection probability increases.The composite image is constructed by overlaying 3 stains: Hoechst (blue), Phalloidin (red), and ConA (green).

concentration of a treatment with antiviral potency against a specific target, up to the point that it does not cause toxicity, interfere with vital cellular functions, or drastically alter cell morphology. Importantly, most of the identified compounds were previously demonstrated to have antiviral activity against SARS-CoV-2, by other drug repurposing studies for COVID-19 based on morphological profiling [63, 38, 22], supporting DEEMD's methodology. On the other hand, if we compare the same plots for the predictions based on the whole-image based model $\mathcal{W}(\phi)$ to those of the DEEMD, we notice that the estimated efficacy scores are scattered randomly among all treatments and lack any form of structure or pattern, Fig 5.4-(c). As mentioned in Section 5.1, the patch-based model $\mathcal{V}(\psi)$ is a hybrid of the other two models in terms of its training. This hybridization also manifests itself in Fig 5.4-(b). We can see that the estimations are slightly structured, not as much as DEEMD's. Implying that the model was not able to extract the informative features and information completely due to its noisy training environment.

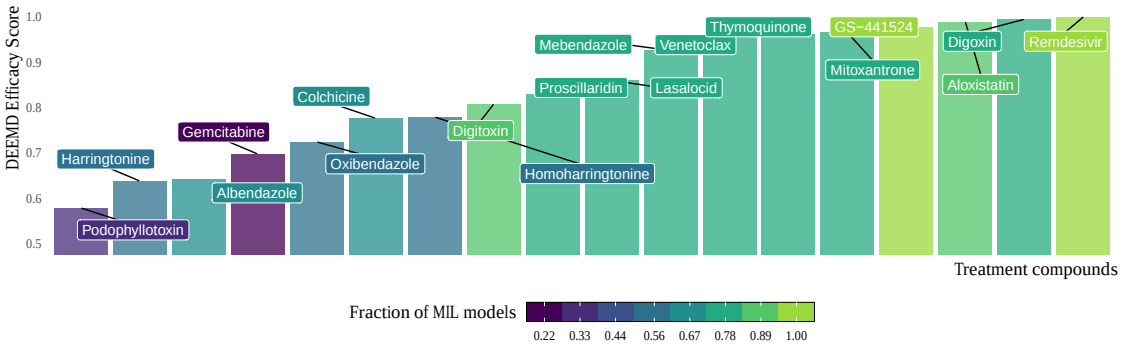
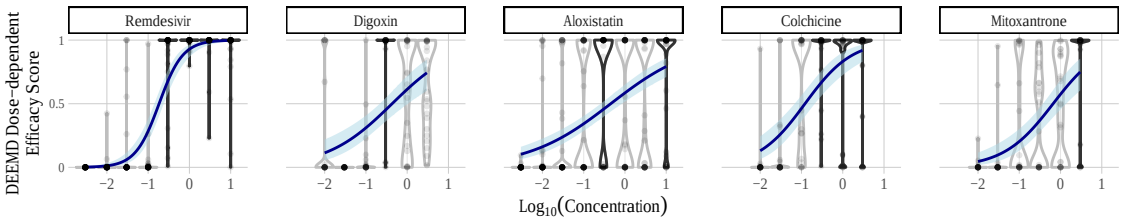


Figure 5.3: Ranked list of DEEMD identified treatments $\mathcal{E}_{\mathcal{M}(\theta)}$ with their DEEMD efficacy scores e_{t_i} on the y-axis. The bars are color-coded based on the fraction of the MIL models that identify it as effective.

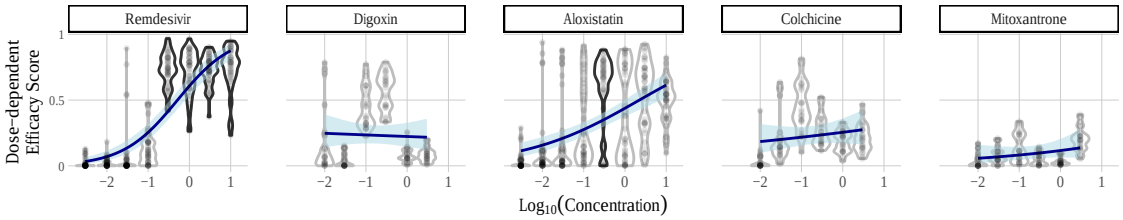
5.4 DEEMD identified treatments are reoccurring in the literature

DEEMD has identified treatments that are reported in the literature to possess therapeutic activity against SARS-CoV-2 [63, 38, 22, 109]. In Fig 5.5 we compare DEEMD efficacy scores to hit-scores reported by Heiser *et al.* [38], which similarly used the RxRx19a dataset. We can see that reoccurring treatments with high DEEMD efficacy scores, namely *Remdesivir*, *GS-441524*, and *Aloxistatin* were also assigned a high hit-score by Heiser *et al.* ($\rho = 0.56, p = 0.02$). In the following, we review DEEMD top ranked identified treatments and briefly discuss their potential mechanisms of action against SARS-CoV-2 infection. In Table 5.1, a complete list of studies for all of the identified treatments are presented. We also compare the DEEMD dose-dependent efficacy score for these compounds to previously reported IC50 values as a measure of validity whenever such data is available; however we acknowledge that IC50 values from *in vitro* studies might not be exactly transferable to clinically relevant concentrations. It should also be noted that due to differences in experimental design, such as cell line timing of treatment and duration of infection, IC50 values from *in vitro* studies can vary widely.

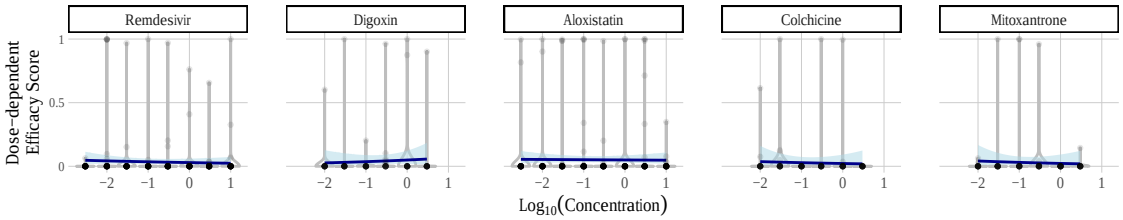
- The most well-known treatment in $\mathcal{E}_{\mathcal{M}(\theta)}$ is *Remdesivir* and its metabolite, *GS-441524*. Previously studied for their antiviral effectiveness against Ebola virus, these compounds target the virus-encoded RNA-dependent RNA polymerase complex needed by RNA viruses to replicate their genome [33]. Multiple studies and clinical trials have found that both compounds are effective against SARS-CoV-2, which led to its emergency approval by the FDA [109, 38, 22, 63] (and refrences within). DEEMD dose-dependent efficacy scores for *Remdesivir* are shown in Fig 5.4-(a). We can see that the DEEMD efficacy score is persistently close to 1 for concentrations in $0.3 - 10$ μM which is consistent with IC50 values reported by other *in vitro* studies [93].



(a) DEEMD dose-dependent efficacy scores based on MIL model $\mathcal{M}(\theta)$ with $k = 2$



(b) Dose-dependent efficacy scores based on the patch-based model $\mathcal{V}(\psi)$



(c) Dose-dependent efficacy scores based on the whole-image based model $\mathcal{W}(\phi)$

Figure 5.4: Estimated efficacy scores for top ranked identified treatments in $\mathcal{E}_{\mathcal{M}(\theta)}$. **(a)** Estimated Dose-dependent efficacy scores for DEEMD top-ranked treatments with $k = 2$. The x and y axes show Log_{10}(Concentration) and dose-dependent efficacy score for each compound. The violins show the estimated distribution of estimated scores for each concentration, along with the data points. A violin is opaque if DEEMD identified that concentration to be effective, i.e. $e_{t_i}^{cj} \geq \zeta$. We used all of the data points, $\mathcal{T}_{t_i}$, to fit a logistic regression to better visualize the trend. **(c)** and **(b)** are similar plots for predictions based on the whole-image based model $\mathcal{W}(\phi)$ and the patch-based model $\mathcal{V}(\psi)$ respectively.

- *Digoxin* is a treatment used for heart disease with a well-established safety profile. Multiple *in vitro* drug repurposing studies for COVID-19 reported its ability to inhibit SARS-CoV-2 replication [93]. The exact mechanism of action for viral inhibition is not identified yet, however, Cho *et al.* [16] hypothesised that *Digoxin* inhibition occurs at the step of viral RNA synthesis. They used multiple FDA-approved treatments, including *Digoxin*, on SARS-CoV-2 infected Vero cells. To understand how the drugs might inhibit SARS-CoV-2, they were administered at three different time points:

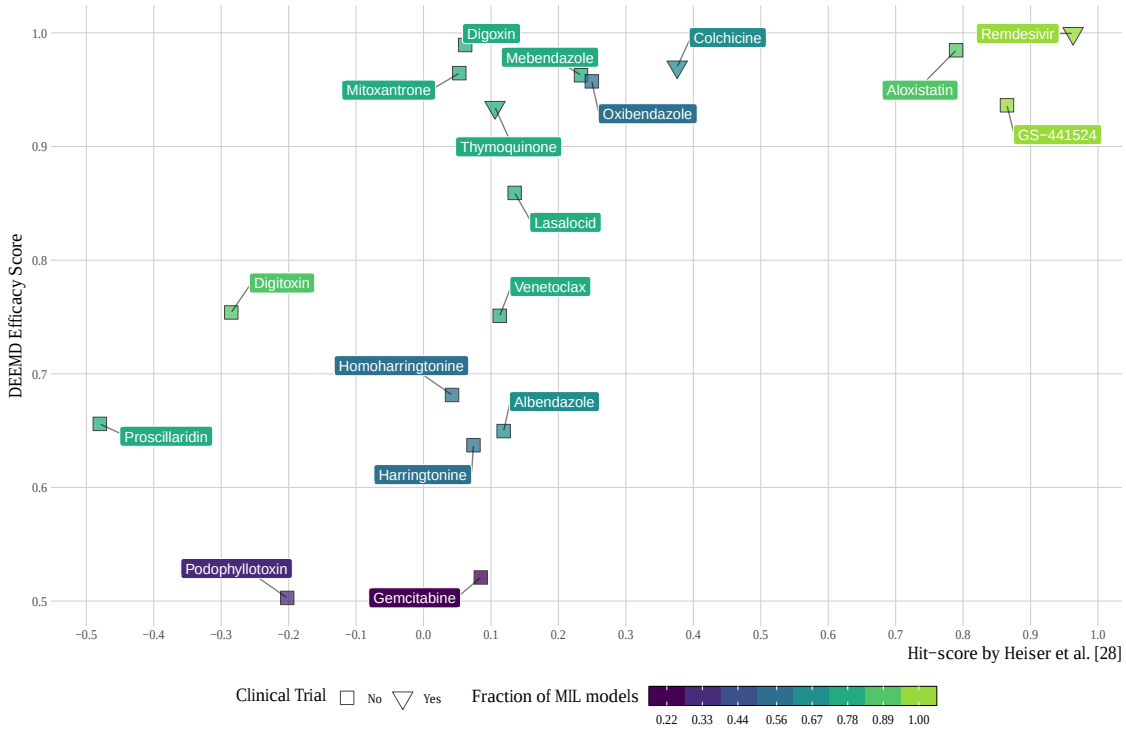


Figure 5.5: Comparison of DEEMD efficacy scores and Heiser *et al.* hit-scores [38] for treatments in $\mathcal{E}_{\mathcal{M}(\theta)}$. Each data point is color-coded based on the fraction of the MIL models that identify it as effective and its shape indicates whether the treatment has gone into clinical trial against COVID-19. The y-axis represents the DEEMD efficacy scores and hit-scores are on the x-axis.

1) prior to infection (prophylactic), 2) at the time of infection (entry), and 3) after the infection (therapeutic). They reported that *Digoxin* showed high efficacy following prophylactic and therapeutic administration but failed to effectively reduce viral replication when administered at the time of infection. However, it should be noted that *Digoxin* exhibited cytotoxicity, as reported by Mirabelli *et al* [63]. Fig 5.4-(a) shows the DEEMD estimated dose-dependent efficacy scores for *Digoxin*; it passed the threshold on 0.3 μM which is aligned with multiple studies that found an *in vitro* IC50 of 0.2 μM. We hypothesize that the DEEMD dose-dependent efficacy scores are not conclusive for higher concentrations because of *Digoxin* reported cytotoxicity.

- Through multiple clinical trials, low doses of *Colchicine* were found to reduce the risk of cardiovascular events in both patients with cardiovascular disease and those patients who had experienced a myocardial infarction [66]. *Colchicine* is an orally administered anti-inflammatory drug, which inhibits polymerization of tubulin and microtubule assembly, and affects a range of cellular immunity pathways [109]. The

clinically demonstrated effects on inflammation and associated illnesses have positioned *Colchicine* as a candidate drug to treat COVID-19. While previous works have presented conflicting conclusions regarding the efficacy of *Colchicine* [24], multiple clinical trials have found it to be effective against SARS-CoV-2 [92]. As shown in Fig 5.4-(a), DEEMD predicted it to be effective at concentrations of $0.3 - 3$ μM.

- *Aloxistatin* (*E*-64*d*) is another potential candidate treatment reported in the literature as an effective agent against SARS-CoV-2. *Aloxistatin* is a membrane-permeable irreversible cysteine-protease inhibitor of calpains and cathepsins. Recent studies have shown that SARS-CoV-2 requires cathepsin L to enter some cell types. *Aloxistatin* can significantly reduce entry of SARS-CoV-2 pseudovirions by inhibiting cathepsin L. [71, 109, 82] (and references within). Fig 5.4-(a) shows DEEMD estimated dose-dependent efficacy scores for *Aloxistatin*. As we can see, the estimated efficacy is higher than cutoff threshold ζ at 0.3 and 10 μM. Olaleye *et al.* used Vero cells to investigate the antiviral activity of several compounds including *Aloxistatin* for which they reported an *in vitro* IC50 of 22 μM [68]. Two additional studies [38] and [22] using morphology-based approaches found that *Aloxistatin* shows strong efficacy without inducing morphological changes to cells.

- DEEMD also identified *Mitoxantrone* to be effective against SARS-CoV-2, as shown in Fig 5.4-(a). Fig 5.8 suggests that there is a consensus between other MIL models on its effectiveness. The cell surface heparan sulfate (HS) is a molecule commonly found on the membrane and on extracellular proteins of cells that assists the endocytosis of many cargos, including SARS-CoV-2 spike. Recently, Zhang *et al.* reported that HS facilitates spike-dependent viral entry of SARS-CoV-2 [111]. They experimented with inhibitor drugs that target the HS-dependent cell entry pathway, and observed that *Mitoxantrone* inhibited viral entry by directly binding to cell surface HS.

- DEEMD also identified three compounds, *Mebendazole*, *Oxibendazole* and *Albendazole* which belong to a large chemical family of *Benzimidazoles* that are used to treat nematode and trematode infections, Fig 5.4-(a) and Fig 5.8. These three compounds specifically have not yet been demonstrated to exhibit activity against SARS-CoV-2 or COVID-19. However, some antiviral activity against other viruses have been found using *Benzimidazole* derivatives [36] and thus based on our analysis, these should be further examined for activity and possible mechanism of action against SARS-CoV-2. It is worth mentioning that these compounds have been suggested in some computational drug repurposing studies to have effectiveness against SARS-CoV-2, but the efficacy of them in animal or human models is yet to be tested [96, 29, 53].

Rank	Compound	Studies
1	Remdesivir	[109, 38, 22, 63, 93, 55, 77, 100, 106, 110, 17]
2	Digoxin	[93, 16]
3	Aloxistatin	[71, 109, 82, 68, 38, 22]
4	Colchicine	[109, 24, 92, 35, 81]
5	Mitoxantrone	[111]
6	Mebendazole	[96, 29, 53]
7	Oxibendazole	[53]
8	GS-441524	[83, 55, 77, 100, 106, 110]
9	Thymoquinone	[2, 28, 87, 102]
10	Lasalocid	[90]
11	Digitoxin	[44, 48, 76]
12	Venetoclax	[31]
13	Homoharringtonine	[106, 13, 17, 42, 109, 41]
14	Proscillaridin	[32]
15	Albendazole	[53]
16	Harringtonine	-
17	Gemcitabine	[106, 113, 41]
18	Podophyllotoxin	[39]

Table 5.1: Ranked list of DEEMD identified treatments, $\mathcal{E}_{\mathcal{M}(\theta)}$, along with drug repurposing or clinical studies that reported to show effectiveness against SARS-CoV-2 or COVID-19.

5.5 DEEMD identified treatments are sensitive to selection of k

Choosing an optimal value for hyper-parameter k is crucial to the performance of MIL models since it effectively controls the amount of noise in the training labels. The hyper-parameter k is the number of top high probability instances in each bag used for loss calculation in each training iteration, 3.4. Tuning this hyper-parameter can be either based on problem-specific domain knowledge or a performance metric on the validation set. For the problem of estimating the treatment efficacy against SARS-CoV-2, the choice of optimal k could be based on the MOI for the infectious agent. However, due to the limited understanding of SARS-CoV-2 and insufficient experimental details about the data generation procedure, we turned to using a performance metric. Using our main objective of identifying effective treatments for tuning hyper-parameters is not feasible because no ground truth data about effective treatments against SARS-CoV-2 exists. Instead, we quantify the hyper-parameter search space by each model's performance on classification of SARS-CoV-2 infected versus non-infected sample images in the validation set. We measured and compared the average precision for MIL models with $k \in \{1, 2, 3, 5, 10, 15, 16, 25, 49\}$ and the MIL model with $k = 2$ showed the best performance. When we compared the dose-dependent efficacy scores for the treatments in $\mathcal{E}_{\mathcal{M}(\theta)}$, with $k = 2$, across different values for

k, Fig 5.8, we observed that in most cases, the fitted logistic regression retained its shape and the inflection points are close together; suggesting a consensus, almost independent of choice of k, between the MIL models on the degree of effectiveness for various treatments.

When comparing DEEMD configurations with different values for k, multiple reoccurring treatments were observed. As shown in Fig 5.5, most of the treatments in $\mathcal{E}_{\mathcal{M}(\theta)}$ have a high recurrence rate in the MIL models, implying their distinguishable ability to suppress viral-induced morphological changes. We noticed that based on the patch-based model $\mathcal{V}(\psi)$ only three treatments were identified to be effective: *Remdesivir*, *GS-441524*, and *Aloxistatin*, which are among the most commonly identified treatments both in COVID-19 drug repurposing studies and multiple MIL models with different k, as shown in Fig 5.5. This observation suggests that for high values of k the amount of noisy labels drastically changes the dynamics of the training. The learned feature space lacks the required sensitivity for downstream analysis of treatment identification, while the model is still capable of accurately classifying the sample images into infected and non-infected classes.

5.6 Using MIL results in less noisy labels in the training compared to the patch-based models

As mentioned in subsection 4.3, the patch-based model $\mathcal{V}(\psi)$ trains the model in the presence of noisy labels. To better understand this model and compare it to the MIL model $\mathcal{M}(\theta)$, it would be insightful to measure how much noise is in the training labels for both models. Each image is split into N patches and there are N_p positive samples and N_n negative samples in the dataset. We define the label noise in the dataset to be the ratio of the samples correctly labeled in the dataset to the total number of samples. Without loss of generality, assume that the fraction of the patches in a sample image that are truly infected has an expected value of λ. Therefore, the expected noise ratio (NR) for the MIL model $\mathcal{M}(\theta)$ and the patch-based model $\mathcal{V}(\psi)$ can be expressed as follows:

$$\mathrm{NR}_{\mathcal{M}(\theta)} = \frac{|\frac{k}{N} - \lambda| N_P}{N_P + N_N}, \quad \mathrm{NR}_{\mathcal{V}(\psi)} = \frac{(1-\lambda) N_P}{N_P + N_N}. \tag{5.1}$$

By defining $r(\lambda, \frac{k}{N})$ to be the log ratio of $\mathrm{NR}_{\mathcal{V}(\psi)}$ to $\mathrm{NR}_{\mathcal{M}(\theta)}$, we can quantitatively analyze these two models behaviour in different configurations for λ and k.

$$r(\lambda, \frac{k}{N}) \triangleq \ln\Big(\frac{\mathrm{NR}_{\mathcal{V}(\psi)}}{\mathrm{NR}_{\mathcal{M}(\theta)}}\Big) = \ln\Big(\frac{1-\lambda}{|\frac{k}{N} - \lambda|}\Big). \tag{5.2}$$

When $r(\lambda, \frac{k}{N}) \geq 0$ the MIL model $\mathcal{M}(\theta)$ has less noisy labels compared to the patch-based model $\mathcal{V}(\psi)$. The landscape of $r(\lambda, \frac{k}{N})$ is visualized in Fig 5.6 for closer inspection. To ensure the numerical stability of $r(\lambda, \frac{k}{N})$ and keep it bounded on the $y = x$ line, a small value ϵ was added to $\mathrm{NR}_{\mathcal{M}(\theta)}$.

One can notice that in almost three forth of the cases, $r(\lambda, \frac{k}{N}) \geq 0$, meaning that the MIL model $\mathcal{M}(\theta)$ has less noisy labels in its training procedure on average, hence, the model converges faster to an optimum point. Moreover, the MIL model $\mathcal{M}(\theta)$ training procedure results in a more efficient training both in terms of computation footprint and extracted features. The model is using only $\frac{k}{N}$ of the training set for updating the weights. This is $\frac{N}{k}$ times faster than using the whole dataset in the patch-based model $\mathcal{V}(\psi)$. And more importantly, the patches that are selected for training the MIL model $\mathcal{M}(\theta)$, would contain more discriminative features since they were top-ranked among all patches in their samples for their informativeness. Finally, the MIL model $\mathcal{M}(\theta)$ has the capability of incorporating domain knowledge about the problem into the training by the choice of hyper-parameter k. By choosing k based on prior or domain knowledge, the MIL model $\mathcal{M}(\theta)$ can iteratively refine its training dataset towards for less noisy labels.

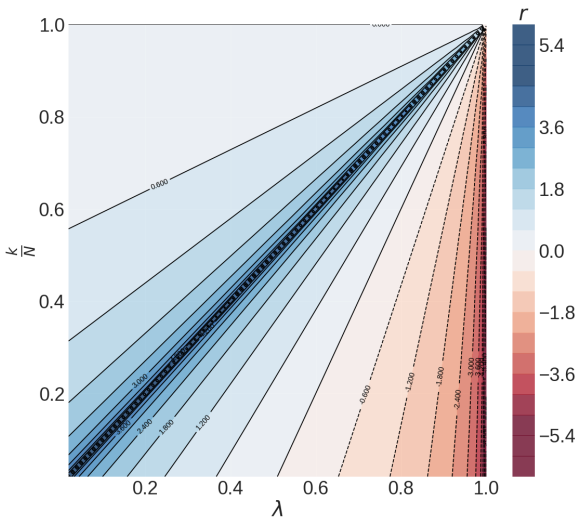


Figure 5.6: Contour plot for $r(\lambda, \frac{k}{N})$, all possible values for λ and $\frac{k}{N}$ are shown on the x and y axis respectively.

5.7 The MIL model identifies predictive patches

Localization of infected regions in a sample image can be done by aggregating all estimated patch-level labels $\mu_i^{p_j}$ to generate an infection map (refer to Section 3.3 for details on the aggregation procedure). These maps can be used to better understand and explain the MIL model $\mathcal{M}(\theta)$ classification predictions as well as providing an annotation for patches that contain SARS-CoV-2 infected cells. Notice that in the calculation of these maps, no form of annotation was used; instead, the MIL model uses sample-level labels to localize the infection. Fig 5.7 shows some representative examples from the untreated test set, overlaid with infection maps predicted by the MIL model $\mathcal{M}(\theta)$. Visual inspection of the infection maps and regions that are predicted with high infection probability, annotated as bright white, suggests that the MIL model $\mathcal{M}(\theta)$ is focused on high intensity stained areas, which we predict to reflect cell death or cytopathic effect, detectable in all channels, except for Hoechst.

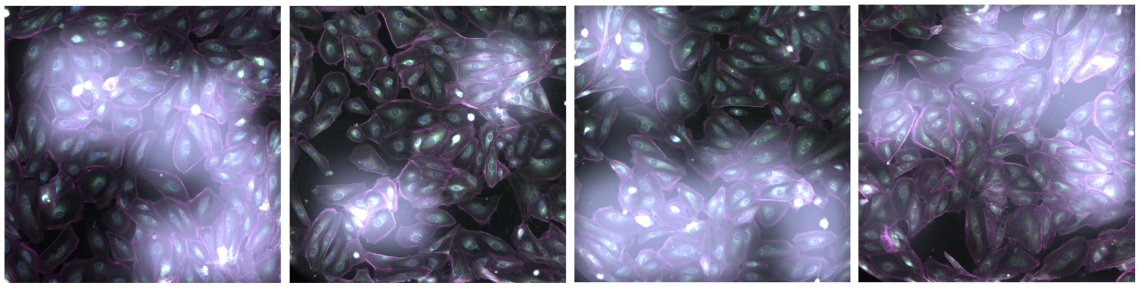


Figure 5.7: Infection map for a diverse set of examples. Representative examples from the untreated test set overlaid with their corresponding infection maps. The intensity of the white color shows the probability of infection. The composite image is constructed by overlaying each stain in a color channel: Hoechst (violet), Syto14 (red), Phalloidin (green), ConA (blue), and WGA (cyan).

5.8 MIL infection maps and cytopathic effects

We further inspected the infection maps generated by the MIL model $\mathcal{M}(\theta)$, Fig 5.7, on the infected sample images from the test set to investigate if we can understand or interpret the patches and features that are triggering the model to identify them as infected. The high infection probability patches are mostly concentrated around high intensity stained features, detectable in all channels, except for Hoechst. These are presumed to be dead or damaged cells or other cytopathic effects (CPE) resulting from SARS-CoV-2 infection. CPE refers to changes in host cell structure as a result of viral infection and SARS-CoV-2 is known to be a cytopathogenic agent [117] for example SARS-CoV-2 is known to cause cell death and to induce syncytia formation (fusion of adjacent cell membranes). CPE can be measured indirectly by using luminescent cell viability assays [80], however, the RxRx19a dataset does not include a specific viability marker, hence CPE can not be quantified on these samples. Nonetheless, some forms of putative CPE are detectable in the highlighted regions in the infection maps. We hypothesize that the brightly stained areas are cells or subcellular structures that are taking up the cell dyes non-specifically due to changes associated with cell death such as loss of membrane integrity, cell shrinkage, and nuclear fragmentation *etc.* [23]. In support of this, we also observed that the cell nucleus count is significantly lower in the infected samples which is also suggestive of cell death. This suggests that the MIL model $\mathcal{M}(\theta)$ has incorporated biologically relevant morphological features into the infection map, as is expected.

5.9 DEEMD is limited by drug toxicity

Drug toxicity refers to a compound's negative side effects on a living cell, as the compound can disrupt crucial cellular functions and pathways to the extent of causing cell death.

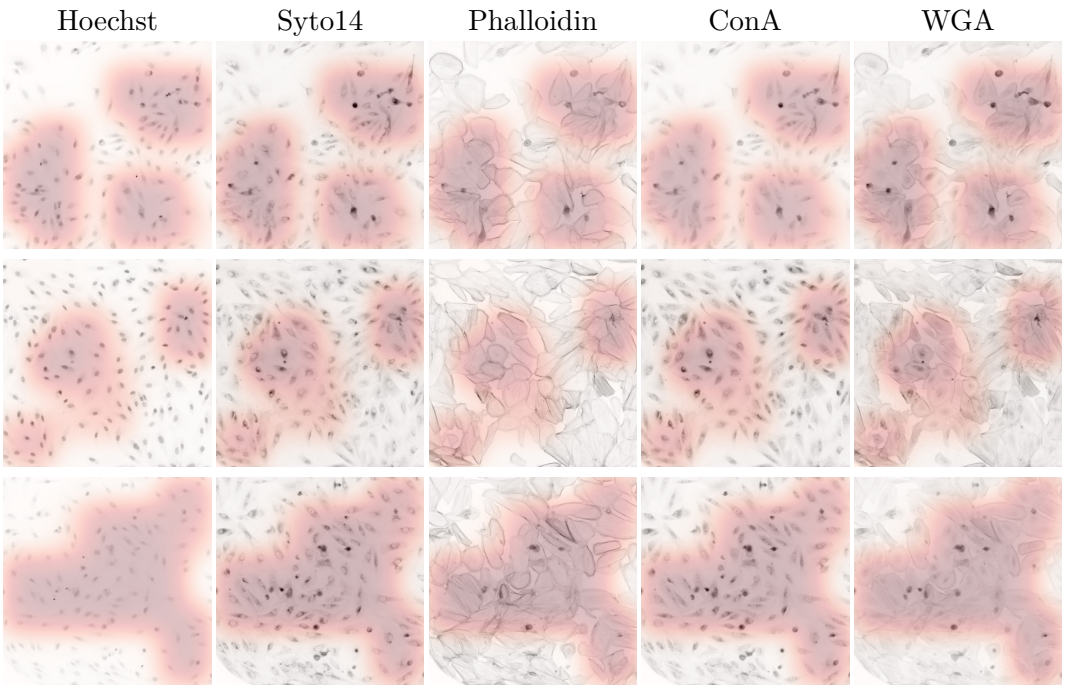


Table 5.2: More examples of the infection maps seperated for each stains. For better visualization, each dye image is color-inverted.

Taking drug toxicity into consideration is essential for optimizing the concentration of the compound needed for optimal efficacy [65]. It is noteworthy to clarify that a low efficacy score $e_{t_i}^{c_j}$ estimated by DEEMD does not necessarily indicate dose or treatment ineffectiveness against SARS-CoV-2. If the treatment has toxic effects or induces other changes in cell morphology, the sample morphology may no longer resemble either the uninfected or SARS-CoV-2 infected class; subsequently, the model's prediction would not be conclusive. Currently, DEEMD only relies on the assumption that if the drug is toxic, the treated cell morphology would not be similar to uninfected cells, thus the model estimates a low efficacy for a toxic compound. A more complex model capable of integrating drug toxicity with cell morphology is required to properly capture the dynamics of the treatment compound, toxicity and effects on cellular morphology, including cell death.

The RxRx19a public dataset does not include images from healthy drug treated cells, which restricts the morphological feature space learned by the model during training. The model can only learn the regions corresponding to the uninfected and SARS-CoV-2 infected cell morphology, and it is unaware of the space structure outside of these regions. Including additional classes into the training dataset would allow the model to learn a wider range of morphological variations and enable it to differentiate between treatments that are ineffective from those that are impacting cell morphology through toxicity or any other mechanisms; thus a better control over the false negative rate would be in place. We noticed clear cases of drug toxicity with multiple compounds in the dataset, where high

concentrations clearly disrupted cell morphology. These sample images lack a detectable signal for the different cell structures and thus we decided to exclude them from the treated test set. Training the model to identify cellular morphology associated with drug toxicity would support the identification of the compound and corresponding concentrations with optimal efficacy and minimal toxicity [65].

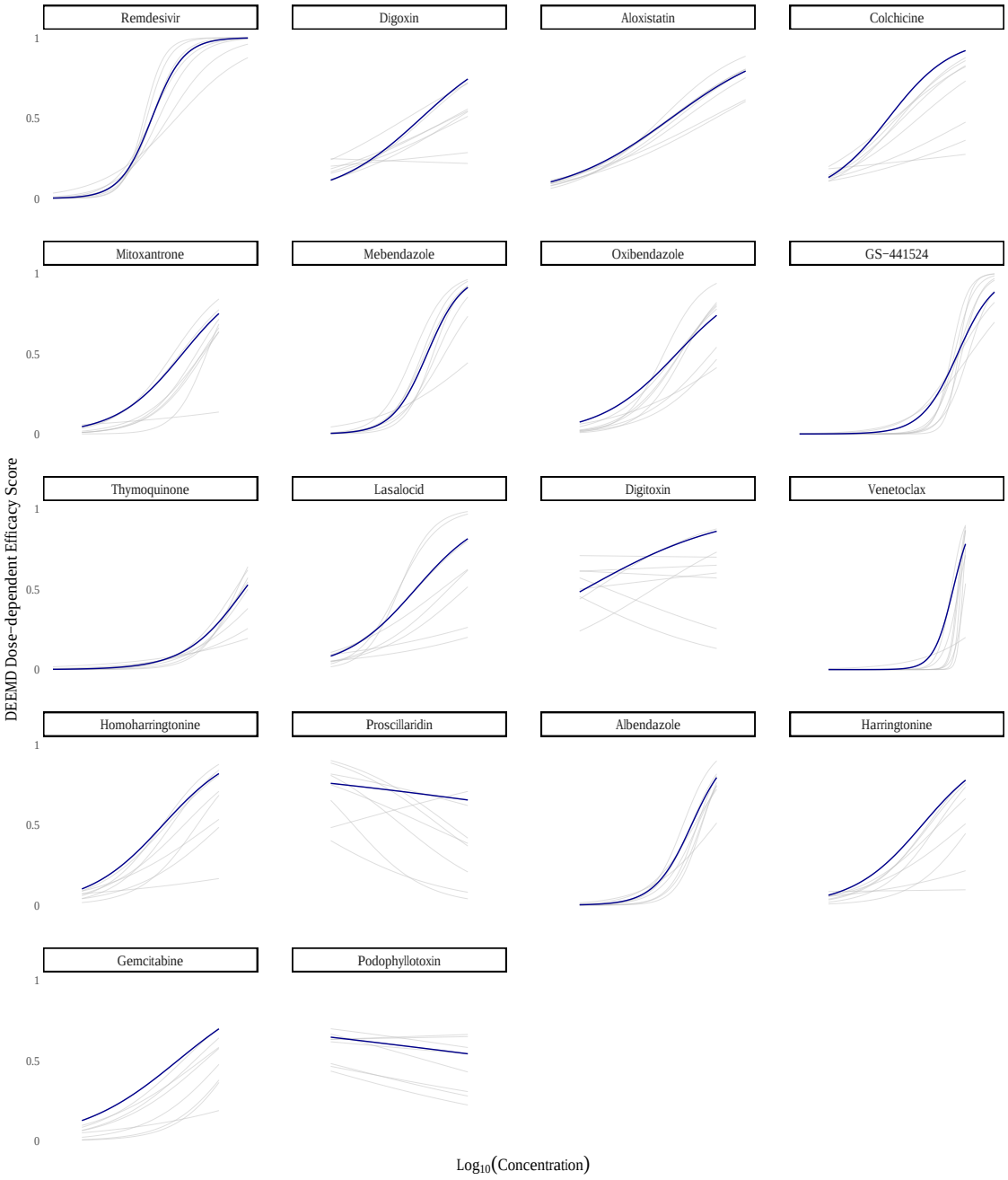


Figure 5.8: DEEMD estimated efficacy scores for identified treatments for different k values. Similar plots to those in Fig 5.4 for treatments in $\mathcal{E}_{\mathcal{M}(\theta)}$, based on multiple choices of k. The logistic curves are fitted based on each value for k. The blue curves represents the MIL model with $k = 2$ which showed the best classification performance on the validation set.

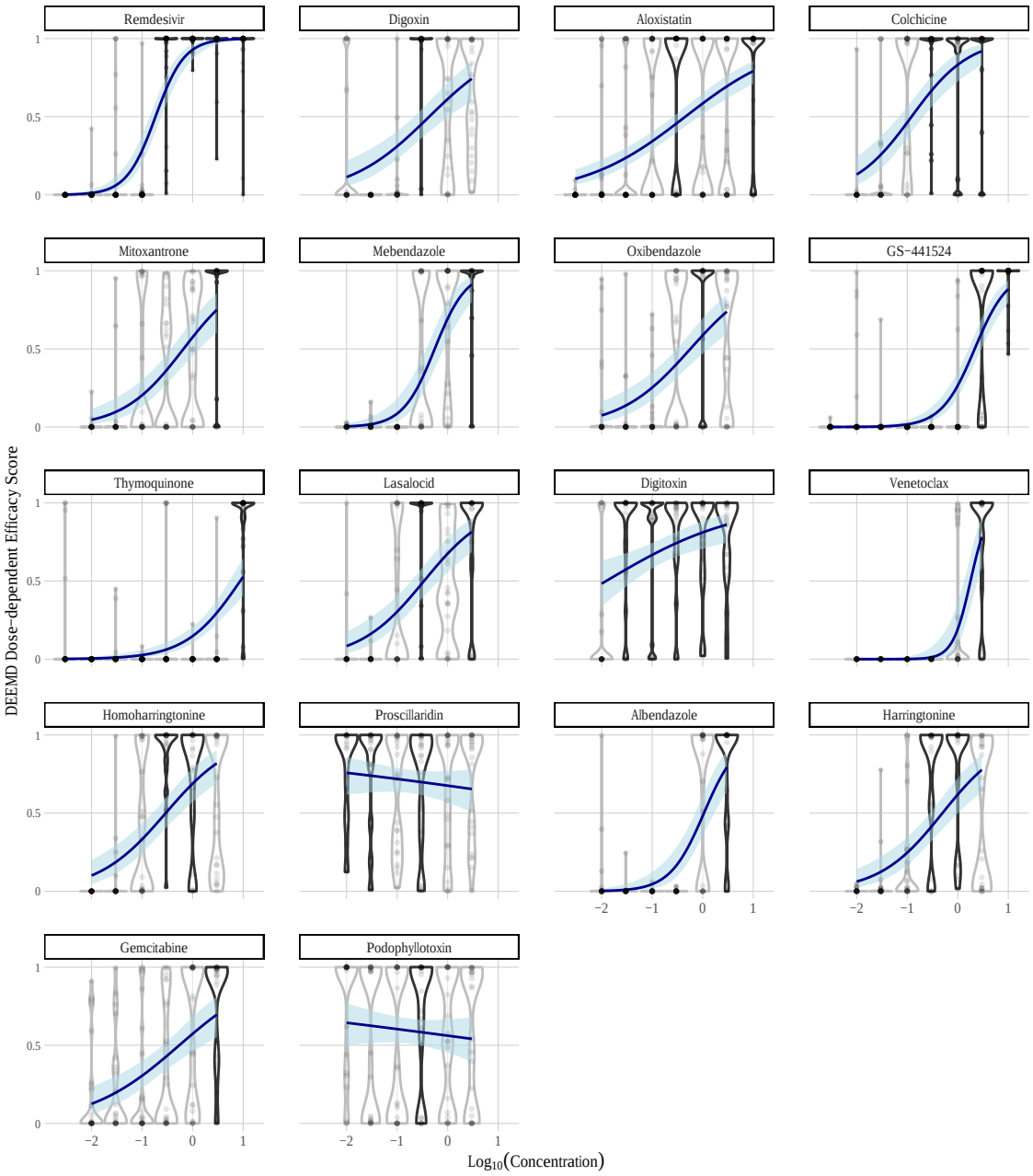


Figure 5.9: DEEMD dose-dependent efficacy scores for all identified treatments in $\mathcal{E}_{\mathcal{M}(\theta)}$. DEEMD dose-dependent efficacy scores in form of violin plots, similar to Fig 5.4. X-axis shows the Log_{10}(Concentration) of each compound and y-axis reflects the estimated efficacy score. The violin plot is opaque if $e_{t_i}^{c_j} > \zeta$ and the blue curve is a logistic regression model fit to each $\mathcal{T}_{t_i}$ along with its shaded 95% confidence interval.

Chapter 6

Conclusion

In this work, we present DEEMD: a pipeline capable of estimating the treatment efficacy of compounds based on morphological analysis of fluorescent-labelled cells. It includes a deep learning model trained within a MIL framework to extract morphological features corresponding to the predicted SARS-CoV-2 infection versus no infection in micro-populations, as well as generating an infection map in a weakly supervised fashion. We compared the performance of the MIL model $\mathcal{M}(\theta)$ to the conventionally trained whole-image based model $\mathcal{W}(\phi)$. As discussed in Chapter 5, both models are capable of accurately distinguishing between images from uninfected and SARS-CoV-2 infected sample images. By integrating a statistical test into the pipeline, DEEMD identifies efficacious compounds that have been reported to have antiviral effectiveness against SARS-CoV-2 using other methods, supporting the performance of the proposed pipeline, whereas the whole-image based model $\mathcal{W}(\phi)$ fails to estimate meaningful predictions.

High-throughput screening of antiviral compounds against a virus in cell-based systems requires the use of commercially available virus-specific antibodies or engineering of fluorescent-tagged virus particles which are not trivial to develop. Although these molecular tools are invaluable, the cost and time they take to develop can be a limiting factor in the initial stages of an outbreak. In that regard, MIL models have an advantage as they may identify infected cell populations without the need for virus-specific molecular tools. The RxRx19a dataset was released at the very early stages of the pandemic, hence it lacks a specific marker for SARS-CoV-2 that could verify if the MIL model $\mathcal{M}(\theta)$ has correctly annotated specific patches of cells within an image as infected. Nonetheless, the MIL model $\mathcal{M}(\theta)$ is able to accurately distinguish sample images from infected versus non-infected cells. Validation of the model on images with a marker for SARS-CoV-2 infection would provide additional confidence to the proposed pipeline and improve the robustness of the MIL model predictions. This can be done using the viral marker as an attention map within the training procedure, or as a new term to the model's loss function.

We only applied DEEMD to the RxRx19a dataset [38] in this work. With the current configuration, DEEMD is not transferable to other datasets, cell lines, or viruses because it

was only trained on the HRCE cells with SARS-CoV-2 infection. A version of DEEMD that is capable of identifying infection-induced morphological changes across multiple viruses and cell lines can be achieved by leveraging a compiled dataset of multiple cell lines and viruses. The training procedure would be the same except for the addition of an auxiliary input that conditions the model based on the cell line and the virus used in the generation of each sample input image.

DEEMD is designed to be adopted and applied to other datasets and problems with few adjustments and modifications. With proper assumptions and data preprocessing, this pipeline can be applied to any fluorescence microscopy datasets, regardless of the stains and image size, for treatment efficacy estimation.

In the future, we plan to generate and apply DEEMD to more comprehensive datasets that include drug-treated, uninfected cells along with specific markers of viral infection that can be used to properly address the shortcomings and limitations of the current version of DEEMD. Having treated uninfected samples would have a great impact on morphological profiling and would bring machine learning-based drug repurposing one step closer to being widely applied to therapeutics.

Bibliography

[1] *Coronavirus Outbreak*, (accessed September 27, 2020).

[2] Ajaz Ahmad, Muneeb U Rehman, Parvaiz Ahmad, and Khalid M Alkharfy. Covid-19 and thymoquinone: Connecting the dots. *Phytotherapy Research*, 2020.

[3] E Susan Amirian and Julie K Levy. Current knowledge about the antivirals remdesivir (gs-5734) and gs-441524 as therapeutic options for coronaviruses. *One Health*, page 100128, 2020.

[4] Jaume Amores. Multiple instance classification: Review, taxonomy and comparative study. *Artificial intelligence*, 201:81–105, 2013.

[5] Stuart Andrews, Ioannis Tsochantaridis, and Thomas Hofmann. Support vector machines for multiple-instance learning. *Advances in neural information processing systems*, 15:577–584, 2002.

[6] Akshay Avula, Krishna Nalleballe, Naureen Narula, Steven Sapozhnikov, Vasuki Dandu, Sudhamshi Toom, Allison Glaser, and Dany Elsayegh. Covid-19 presenting as stroke. *Brain, behavior, and immunity*, 2020.

[7] Sanghamitra Bandyopadhyay, Dip Ghosh, Ramkrishna Mitra, and Zhongming Zhao. Mbstar: multiple instance learning for predicting specific functional binding sites in microrna targets. *Scientific reports*, 5(1):1–12, 2015.

[8] Mark-Anthony Bray, Shantanu Singh, Han Han, Chadwick T Davis, Blake Borgeson, Cathy Hartland, Maria Kost-Alimova, Sigrun M Gustafsdottir, Christopher C Gibson, and Anne E Carpenter. Cell painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nature protocols*, 11(9):1757, 2016.

[9] Volker Brinkmann, Andreas Billich, Thomas Baumruker, Peter Heining, Robert Schmouder, Gordon Francis, Shreeram Aradhye, and Pascale Burtin. Fingolimod (fty720): discovery and development of an oral drug to treat multiple sclerosis. *Nature reviews Drug discovery*, 9(11):883–897, 2010.

[10] Gabriele Campanella, Matthew G Hanna, Luke Geneslaw, Allen Miraflor, Vitor Werneck Krauss Silva, Klaus J Busam, Edi Brogi, Victor E Reuter, David S Klimstra, and Thomas J Fuchs. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature medicine*, 25(8):1301–1309, 2019.

[11] Marc-André Carbonneau, Veronika Cheplygina, Eric Granger, and Ghyslain Gagnon. Multiple instance learning: A survey of problem characteristics and applications. *Pattern Recognition*, 77:329–353, 2018.

[12] Rudragouda Channappanavar and Stanley Perlman. Pathogenic human coronavirus infections: causes and consequences of cytokine storm and immunopathology. In *Seminars in immunopathology*, volume 39, pages 529–539. Springer, 2017.

[13] Yanwen Chen, Travis Lear, John Evankovich, Mads Larsen, Bo Lin, Irene Alfaras, Jason Kennerdell, Laura Salminen, Daniel Camarco, Karina Lockwood, et al. A high throughput screen for tmprss2 expression identifies fda-approved and clinically advanced compounds that can limit sars-cov-2 entry. 2020.

[14] Vincent CC Cheng, Susanna KP Lau, Patrick CY Woo, and Kwok Yung Yuen. Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. *Clinical microbiology reviews*, 20(4):660–694, 2007.

[15] Veronika Cheplygina, David MJ Tax, and Marco Loog. Multiple instance learning with bag dissimilarities. *Pattern recognition*, 48(1):264–275, 2015.

[16] Junhyung Cho, Young Jae Lee, Je Hyoung Kim, Sang il Kim, Sung Soon Kim, Byeong-Sun Choi, and Jang-Hoon Choi. Antiviral activity of digoxin and ouabain against sars-cov-2 infection and its implication for covid-19. *Scientific reports*, 10(1):1–8, 2020.

[17] Ka-Tim Choy, Alvina Yin-Lam Wong, Prathanporn Kaewpreedee, Sin-Fun Sia, Dongdong Chen, Kenrie Pui Yan Hui, Daniel Ka Wing Chu, Michael Chi Wai Chan, Peter Pak-Hang Cheung, Xuhui Huang, et al. Remdesivir, lopinavir, emetine, and homoharringtonine inhibit sars-cov-2 replication in vitro. *Antiviral research*, page 104786, 2020.

[18] Mirko Cortese, Ji-Young Lee, Berati Cerikan, Christopher J Neufeldt, Viola MJ Oorschot, Sebastian Köhrer, Julian Hennies, Nicole L Schieber, Paolo Ronchi, Giulia Mizzon, et al. Integrative imaging reveals sars-cov-2-induced reshaping of subcellular morphologies. *Cell host & microbe*, 28(6):853–866, 2020.

[19] Nicolas Coudray, Paolo Santiago Ocampo, Theodore Sakellaropoulos, Navneet Narula, Matija Snuderl, David Fenyö, Andre L Moreira, Narges Razavian, and Aristotelis Tsirigos. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature medicine*, 24(10):1559–1567, 2018.

[20] Michael J Cox, Steffen Jaensch, Jelle Van de Waeter, Laure Cougnaud, Daan Seynaeve, Soulaiman Benalla, Seong Joo Koo, Ilse Van Den Wyngaert, Jean-Marc Neefs, Dmitry Malkov, et al. Tales of 1,008 small molecules: phenomic profiling through live-cell imaging in a panel of reporter cell lines. *Scientific reports*, 10(1):1–14, 2020.

[21] Jeff Craven. *COVID-19 vaccine tracker*, (accessed September 27, 2020).

[22] Michael F Cuccarese, Berton A Earnshaw, Katie Heiser, Ben Fogelson, Chadwick T Davis, Peter F McLean, Hannah B Gordon, Kathleen-Rose Skelly, Fiona L Weathersby, Vlad Rodic, et al. Functional immune mapping with deep-learning enabled phenomics applied to immunomodulatory and covid-19 drug discovery. *bioRxiv*, 2020.

[23] Brian S Cummings and Rick G Schnellmann. Measurement of cell death in mammalian cells. *Current protocols in pharmacology*, 25(1):12–8, 2004.

[24] Medine Cumhur Cure, Adem Kucuk, and Erkan Cure. Colchicine may not be effective in COVID-19 infection; it may even be harmful? *Clinical Rheumatology*, 39(7):2101–2102, May 2020.

[25] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[26] Thomas G Dietterich, Richard H Lathrop, and Tomás Lozano-Pérez. Solving the multiple instance problem with axis-parallel rectangles. *Artificial intelligence*, 89(1-2):31–71, 1997.

[27] Wilfrid J Dixon and Alexander M Mood. The statistical sign test. *Journal of the American Statistical Association*, 41(236):557–566, 1946.

[28] Sawsan Elgohary, Aya A Elkhodiry, Nada S Amin, Ulrike Stein, and Hend M El Tayebi. Thymoquinone: a tie-breaker in sars-cov2-infected cancer patients? *Cells*, 10(2):302, 2021.

[29] Ayman Farag, Ping Wang, Mahmoud Ahmed, and Hesham Sadek. Identification of fda approved drugs targeting covid-19 virus by structure-based drug repositioning. *ChemRxiv, doi: doi*, 10, 2020.

[30] James Richard Foulds and Eibe Frank. A review of multi-instance learning assumptions. 2010.

[31] Moritz Fürstenau, Petra Langerbeins, Nisha De Silva, Anna Maria Fink, Sandra Robrecht, Julia von Tresckow, Florian Simon, Karin Hohloch, Jolanda Droogendijk, Marjolein van der Klift, et al. Covid-19 among fit patients with cll treated with venetoclax-based combinations. *Leukemia*, 34(8):2225–2229, 2020.

[32] Kaifu Gao, Duc Duy Nguyen, Jiahui Chen, Rui Wang, and Guo-Wei Wei. Repositioning of 8565 existing drugs for covid-19. *arXiv preprint arXiv:2005.10028*, 2020.

[33] Calvin J Gordon, Egor P Tchesnokov, Joy Y Feng, Danielle P Porter, and Matthias Götte. The antiviral compound remdesivir potently inhibits rna-dependent rna polymerase from middle east respiratory syndrome coronavirus. *Journal of Biological Chemistry*, 295(15):4773–4779, 2020.

[34] Aakriti Gupta, Mahesh V Madhavan, Kartik Sehgal, Nandini Nair, Shiwani Mahajan, Tejasav S Sehrawat, Behnood Bikdeli, Neha Ahluwalia, John C Ausiello, Elaine Y Wan, et al. Extrapulmonary manifestations of covid-19. *Nature medicine*, 26(7):1017–1032, 2020.

[35] Timotius Ivan Hariyanto, Devina Adella Halim, Claudia Jodhinata, Theo Audi Yanto, and Andree Kurniawan. Colchicine treatment can improve outcomes of coronavirus disease 2019 (covid-19): A systematic review and meta-analysis. *Clinical and Experimental Pharmacology and Physiology*, 2021.

[36] Julie R. Harris and Peter J. Hotez. 276 - intestinal nematodes. In Sarah S. Long, Charles G. Prober, and Marc Fischer, editors, *Principles and Practice of Pediatric Infectious Diseases (Fifth Edition)*, pages 1373–1381.e3. Elsevier, fifth edition edition, 2018.

[37] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[38] Katie Heiser, Peter F McLean, Chadwick T Davis, Ben Fogelson, Hannah B Gordon, Pamela Jacobson, Brett L Hurst, Ben J Miller, Ronald W Alfa, Berton A Earnshaw, et al. Identification of potential treatments for covid-19 through artificial intelligence-enabled phenomic analysis of human cells infected with sars-cov-2. *bioRxiv*, 2020.

[39] Andreas Hensel, Rudolf Bauer, Michael Heinrich, Verena Spiegler, Oliver Kayser, Georg Hempel, and Karin Kraft. Challenges at the time of covid-19: Opportunities and innovations in antivirals from nature. *Planta medica*, 86(10):659, 2020.

[40] Le Hou, Dimitris Samaras, Tahsin M Kurc, Yi Gao, James E Davis, and Joel H Saltz. Patch-based convolutional neural network for whole slide tissue image classification. In *Proceedings of the ieee conference on computer vision and pattern recognition*, pages 2424–2433, 2016.

[41] Aleksandr Ianevski, Rouan Yao, Svetlana Biza, Eva Zusinaite, Andres Mannik, Gaily Kivi, Anu Planken, Kristiina Kurg, Eva-Maria Tombak, Mart Ustav, et al. Identification of novel antiviral drug combinations in vitro and tracking their development. *bioRxiv*, 2020.

[42] Aleksandr Ianevski, Rouan Yao, Mona Høysæter Fenstad, Svetlana Biza, Eva Zusinaite, Tuuli Reisberg, Hilde Lysvand, Kirsti Løseth, Veslemøy Malm Landsem, Janne Fossum Malmring, et al. Potential antiviral options against sars-cov-2 infection. *Viruses*, 12(6):642, 2020.

[43] Maximilian Ilse, Jakub M Tomczak, and Max Welling. Attention-based deep multiple instance learning. *arXiv preprint arXiv:1802.04712*, 2018.

[44] Sangeun Jeon, Meehyun Ko, Jihye Lee, Inhee Choi, Soo Young Byun, Soonju Park, David Shum, and Seungtaek Kim. Identification of antiviral drug candidates against sars-cov-2 from fda-approved drugs. *Antimicrobial Agents and Chemotherapy*, 2020.

[45] Jeffrey S Kahn and Kenneth McIntosh. History and recent advances in coronavirus discovery. *The Pediatric infectious disease journal*, 24(11):S223–S227, 2005.

[46] Gary King and Langche Zeng. Logistic regression in rare events data. *Political analysis*, 9(2):137–163, 2001.

[47] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[48] Meehyun Ko, Sangeun Jeon, Wang-Shick Ryu, and Seungtaek Kim. Comparative analysis of antiviral efficacy of fda-approved drugs against sars-cov-2 in human lung cells. *Journal of medical virology*, 93(3):1403–1408, 2021.

[49] Adriano Nunes Kochi, Ana Paula Tagliari, Giovanni Battista Forleo, Gaetano Michele Fassini, and Claudio Tondo. Cardiac and arrhythmic complications in patients with covid-19. *Journal of Cardiovascular Electrophysiology*, 31(5):1003–1008, 2020.

[50] Oren Z Kraus, Jimmy Lei Ba, and Brendan J Frey. Classifying and segmenting microscopy images with deep multiple instance learning. *Bioinformatics*, 32(12):i52–i59, 2016.

[51] Maxim Kuleshov, Daniel JB Clarke, Eryk Kropiwnicki, Kathleen Jagodnik, Alon Bartal, John Erol Evangelista, Abigail Zhou, Laura B Ferguson, Alexander Lachmann, and Avi Ma'ayan. The covid-19 gene and drug set library. 2020.

[52] Michael MC Lai and David Cavanagh. The molecular biology of coronaviruses. In *Advances in virus research*, volume 48, pages 1–100. Elsevier, 1997.

[53] Jeffrey N Law, Nure Tasnina, Meghana Kshirsagar, Judith Klein-Seetharaman, Mark Crovella, Padmavathy Rajagopalan, Simon Kasif, and TM Murali. Identifying human interactors of sars-cov-2 proteins and drug targets for covid-19 using network-based label propagation. *arXiv preprint arXiv:2006.01968*, 2020.

[54] T Thanh Le, Zacharias Andreadakis, Arun Kumar, R Gomez Roman, Stig Tollefsen, Melanie Saville, and Stephen Mayhew. The covid-19 vaccine development landscape. *Nat Rev Drug Discov*, 19(5):305–306, 2020.

[55] Yingjun Li, Liu Cao, Ge Li, Feng Cong, Yunfeng Li, Jing Sun, Yinzhu Luo, Guijiang Chen, Guanguan Li, Ping Wang, et al. Remdesivir metabolite gs-441524 effectively inhibits sars-cov-2 infection in mouse models. *Journal of medicinal chemistry*, 2021.

[56] Lu Lin, Xiayang Jiang, Zhenling Zhang, Siwen Huang, Zhenyi Zhang, Zhaoxiong Fang, Zhiqiang Gu, Liangqing Gao, Honggang Shi, Lei Mai, et al. Gastrointestinal symptoms of 95 cases with sars-cov-2 infection. *Gut*, 69(6):997–1001, 2020.

[57] Guoqing Liu, Jianxin Wu, and Zhi-Hua Zhou. Key instance detection in multi-instance learning. 2012.

[58] Rory KM Long, Kathleen P Moriarty, Ben Cardoen, Guang Gao, A Wayne Vogl, Francois Jean, Ghassan Hamarneh, and Ivan R Nabi. Super resolution microscopy and deep learning identify zika virus reorganization of the endoplasmic reticulum. *Scientific Reports*, 10(1):1–18, 2020.

[59] Elisabeth Mahase. Covid-19: Russia approves vaccine without large scale testing or published results. *BMJ: British Medical Journal (Online)*, 370, 2020.

[60] Oded Maron and Tomás Lozano-Pérez. A framework for multiple-instance learning. In *Advances in neural information processing systems*, pages 570–576, 1998.

[61] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[62] Claire McQuin, Allen Goodman, Vasiliy Chernyshev, Lee Kamentsky, Beth A Cimini, Kyle W Karhohs, Minh Doan, Liya Ding, Susanne M Rafelski, Derek Thirstrup, et al. Cellprofiler 3.0: Next-generation image processing for biology. *PLoS biology*, 16(7):e2005970, 2018.

[63] Carmen Mirabelli, Jesse W Wotring, Charles J Zhang, Sean M McCarty, Reid Fursmidt, Namrata S Kadambi, Anya T Amin, Teresa R O'Meara, Carla D Pretto-Kernahan, Jason R Spence, et al. Morphological cell profiling of sars-cov-2 infection identifies drug repurposing candidates for covid-19. *bioRxiv*, 2020.

[64] SH Myint. Human coronaviruses: a brief review. *Reviews in Medical Virology*, 4(1):35, 1994.

[65] Jeremy K Nicholson, John Connelly, John C Lindon, and Elaine Holmes. Metabonomics: a platform for studying drug toxicity and gene function. *Nature reviews Drug discovery*, 1(2):153–161, 2002.

[66] Stefan M. Nidorf, Aernoud T.L. Fiolet, Arend Mosterd, John W. Eikelboom, Astrid Schut, Tjerk S.J. Opstal, Salem H.K. The, Xiao-Fang Xu, Mark A. Ireland, Timo Lenderink, Donald Latchem, Pieter Hoogslag, Anastazia Jerzewski, Peter Nierop, Alan Whelan, Randall Hendriks, Henk Swart, Jeroen Schaap, Aaf F.M. Kuijper, Maarten W.J. van Hessen, Pradyot Saklani, Isabel Tan, Angus G. Thompson, Allison Morton, Chris Judkins, Willem A. Bax, Maurits Dirksen, Marco Alings, Graeme J. Hankey, Charley A. Budgeon, Jan G.P. Tijssen, Jan H. Cornel, and Peter L. Thompson. Colchicine in patients with chronic coronary disease. *New England Journal of Medicine*, 383(19):1838–1847, 2020.

[67] National Institutes of Health (NIH). *Therapeutic Management of Patients with COVID-19*, (accessed October 27, 2020).

[68] Omonike A Olaleye, Manvir Kaur, Collins C Onyenaka, and Tolulope O Adebusuyi. Discovery of clioquinol and analogues as novel inhibitors of severe acute respiratory syndrome coronavirus 2 infection, ace2 and ace2-spike protein interaction in vitro. *bioRxiv*, 2020.

[69] World Health Organization. *COVID-19 vaccines*, (accessed March 30, 2020).

[70] World Health Organization. *Situation reports*, (accessed September 27, 2020).

[71] Xiuyuan Ou, Yan Liu, Xiaobo Lei, Pei Li, Dan Mi, Lili Ren, Li Guo, Ruixuan Guo, Ting Chen, Jiaxin Hu, et al. Characterization of spike glycoprotein of sars-cov-2 on virus entry and its immune cross-reactivity with sars-cov. *Nature communications*, 11(1):1–12, 2020.

[72] Nikolaos Pappas and Andrei Popescu-Belis. Explicit document modeling through weighted multiple-instance learning. *Journal of Artificial Intelligence Research*, 58:591–626, 2017.

[73] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017.

[74] Rainer Pepperkok and Jan Ellenberg. High-throughput fluorescence microscopy for systems biology. *Nature reviews Molecular cell biology*, 7(9):690–696, 2006.

[75] Thai-Hoang Pham, Yue Qiu, Jucheng Zeng, Lei Xie, and Ping Zhang. A deep learning framework for high-throughput mechanism-driven phenotype compound screening and its application to covid-19 drug repurposing. *Nature Machine Intelligence*, pages 1–11, 2021.

[76] Bette S Pollard, Jorge C Blanco, and John R Pollard. Classical drug digitoxin inhibits influenza cytokine storm, with implications for covid-19 therapy. *in vivo*, 34(6):3723–3730, 2020.

[77] Andrea J Pruijssers, Amelia S George, Alexandra Schäfer, Sarah R Leist, Lisa E Gralinksi, Kenneth H Dinnon III, Boyd L Yount, Maria L Agostini, Laura J Stevens, James D Chappell, et al. Remdesivir inhibits sars-cov-2 in human lung cells and chimeric sars-cov expressing the sars-cov-2 rna polymerase in mice. *Cell reports*, 32(3):107940, 2020.

[78] Sudeep Pushpakom, Francesco Iorio, Patrick A Eyers, K Jane Escott, Shirley Hopper, Andrew Wells, Andrew Doig, Tim Guilliams, Joanna Latimer, Christine McNamee, et al. Drug repurposing: progress, challenges and recommendations. *Nature reviews Drug discovery*, 18(1):41–58, 2019.

[79] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013.

[80] Laura Riva, Shuofeng Yuan, Xin Yin, Laura Martin-Sancho, Naoko Matsunaga, Sebastian Burgstaller, Lars Pache, Paul De Jesus, Mitchell V Hull, Max Chang, et al. A large-scale drug repositioning survey for sars-cov-2 antivirals. *BioRxiv*, 2020.

[81] Naomi Schlesinger, Bonnie L Firestein, and Luigi Brunetti. Colchicine in covid-19: an old drug, new use. *Current Pharmacology Reports*, 6(4):137–145, 2020.

[82] Jian Shang, Yushun Wan, Chuming Luo, Gang Ye, Qibin Geng, Ashley Auerbach, and Fang Li. Cell entry mechanisms of sars-cov-2. *Proceedings of the National Academy of Sciences*, 117(21):11727–11734, 2020.

[83] Yuejun Shi, Lei Shuai, Zhiyuan Wen, Chong Wang, Yan Yuanyuan, Zhe Jiao, Fenglin Guo, Zhen F Fu, Chen Huanchun, Zhigao Bu, et al. The preclinical inhibitor gs441524 in combination with gc376 efficaciously inhibited the proliferation of sars-cov-2 in the mouse respiratory tract. *Emerging Microbes & Infections*, pages 1–43, 2021.

[84] Jaak Simm, Günter Klambauer, Adam Arany, Marvin Steijaert, Jörg Kurt Wegner, Emmanuel Gustin, Vladimir Chupakhin, Yolanda T Chong, Jorge Vialard, Peter Buijnsters, et al. Repurposing high-throughput image assays enables biological activity prediction for drug discovery. *Cell chemical biology*, 25(5):611–618, 2018.

[85] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[86] Tanu Singhal. A review of coronavirus disease-2019 (covid-19). *The Indian Journal of Pediatrics*, pages 1–6, 2020.

[87] Andrei P Sommer, Horst-Dieter Försterling, and Kurt G Naber. Thymoquinone: shield and sword against sars-cov-2. *Precis Nanomed*, 3(3):541–548, 2020.

[88] V Starkuviene and R Pepperkok. The potential of high-content high-throughput microscopy in drug discovery. *British journal of pharmacology*, 152(1):62–71, 2007.

[89] Aravind Subramanian, Rajiv Narayan, Steven M Corsello, David D Peck, Ted E Natoli, Xiaodong Lu, Joshua Gould, John F Davis, Andrew A Tubelli, Jacob K Asiedu, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell*, 171(6):1437–1452, 2017.

[90] Esben B Svenningsen, Jacob Thyrsted, Julia Blay-Cadanet, Han Liu, Shaoquan Lin, Jaime Moyano-Villameriel, David Olagnier, Manja Idorn, Søren R Paludan, Christian K Holm, et al. Ionophore antibiotic x-206 is a potent inhibitor of sars-cov-2 infection in vitro. *Antiviral research*, 185:104988, 2021.

[91] Xiaolu Tang, Changcheng Wu, Xiang Li, Yuhe Song, Xinmin Yao, Xinkai Wu, Yuange Duan, Hong Zhang, Yirong Wang, Zhaohui Qian, et al. On the origin and continuing evolution of sars-cov-2. *National Science Review*, 2020.

[92] Jean-Claude Tardif, Nadia Bouabdallaoui, Philippe L L'Allier, Daniel Gaudet, Binita Shah, Michael H Pillinger, Jose Lopez-Sendon, Protasio da Luz, Lucie Verret, Sylvia Audet, et al. Efficacy of colchicine in non-hospitalized patients with covid-19. *Medrxiv*, 2021.

[93] Philip L Tzou, Kaiming Tao, Janin Nouhin, Soo-Yon Rhee, Benjamin D Hu, Shruti Pai, Neil Parkin, and Robert W Shafer. Coronavirus antiviral research database (cov-rdb): an online database designed to facilitate comparisons between candidate anti-coronavirus compounds. *Viruses*, 12(9):1006, 2020.

[94] Fang Wan, Chang Liu, Wei Ke, Xiangyang Ji, Jianbin Jiao, and Qixiang Ye. C-mil: Continuation multiple instance learning for weakly supervised object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2199–2208, 2019.

[95] Chen Wang, Peter W Horby, Frederick G Hayden, and George F Gao. A novel coronavirus outbreak of global health concern. *The Lancet*, 395(10223):470–473, 2020.

[96] Zhihan Wang, Kai Guo, Pan Gao, Qinqin Pu, Min Wu, Changlong Li, and Junguk Hur. Identification of repurposable drugs and adverse drug reactions for various courses of covid-19 based on single-cell rna sequencing data. *ArXiv*, 2020.

[97] Mathias J Wawer, Kejie Li, Sigrun M Gustafsdottir, Vebjorn Ljosa, Nicole E Bodycombe, Melissa A Marton, Katherine L Sokolnicki, Mark-Anthony Bray, Melissa M Kemp, Ellen Winchester, et al. Toward performance-diverse small-molecule libraries for cell-based phenotypic screening using multiplexed high-dimensional profiling. *Proceedings of the National Academy of Sciences*, 111(30):10911–10916, 2014.

[98] Frank Wilcoxon. Individual comparisons by ranking methods. In *Breakthroughs in statistics*, pages 196–202. Springer, 1992.

[99] Fei Xiao, Meiwen Tang, Xiaobin Zheng, Ye Liu, Xiaofeng Li, and Hong Shan. Evidence for gastrointestinal infection of sars-cov-2. *Gastroenterology*, 158(6):1831–1833, 2020.

[100] Xuping Xie, Antonio E Muruato, Xianwen Zhang, Kumari G Lokugamage, Camila R Fontes-Garfias, Jing Zou, Jianying Liu, Ping Ren, Mini Balakrishnan, Tomas Cihlar, et al. A nanoluciferase sars-cov-2 for rapid neutralization testing and screening of anti-infective drugs for covid-19. *Nature communications*, 11(1):1–11, 2020.

[101] Gang Xu, Zhigang Song, Zhuo Sun, Calvin Ku, Zhe Yang, Cancheng Liu, Shuhao Wang, Jianpeng Ma, and Wei Xu. Camel: A weakly supervised learning framework for histopathology image segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 10682–10691, 2019.

[102] Huan Xu, Bing Liu, Zhen Xiao, Meiling Zhou, Lin Ge, Fan Jia, Yanling Liu, Hongshan Jin, Xiuliang Zhu, Jian Gao, et al. Computational and experimental studies reveal that thymoquinone blocks the entry of coronaviruses into in vitro cells. *Infectious diseases and therapy*, pages 1–12, 2021.

[103] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. Show, attend and tell: Neural image caption generation with visual attention. In *International conference on machine learning*, pages 2048–2057. PMLR, 2015.

[104] Xin Xu and Eibe Frank. Logistic regression and boosting for labeled bags of instances. In *Pacific-Asia conference on knowledge discovery and data mining*, pages 272–281. Springer, 2004.

[105] Yongluan Yan, Xinggang Wang, Xiaojie Guo, Jiemin Fang, Wenyu Liu, and Junzhou Huang. Deep multi-instance learning with dynamic pooling. In *Asian Conference on Machine Learning*, pages 662–677, 2018.

[106] Cheng-Wei Yang, Tzu-Ting Peng, Hsing-Yu Hsu, Yue-Zhi Lee, Szu-Huei Wu, Wen-Hsing Lin, Yi-Yu Ke, Tsu-An Hsu, Teng-Kuang Yeh, Wen-Zheng Huang, et al. Repurposing old drugs as antiviral agents for coronaviruses. *Biomedical journal*, 43(4):368–374, 2020.

[107] Jiawen Yao, Xinliang Zhu, Jitendra Jonnagaddala, Nicholas Hawkins, and Junzhou Huang. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. *Medical Image Analysis*, 65:101789, 2020.

[108] EL-Manzalawy Yasser, Drena Dobbs, and Vasant Honavar. Predicting mhc-ii binding affinity using multiple instance regression. *Ieee/acm transactions on computational biology and bioinformatics*, 8(4):1067–1079, 2010.

[109] Hassan Yousefi, Ladan Mashouri, Samuel C Okpechi, Nikhilesh Alahari, and Suresh K Alahari. Repurposing existing drugs for the treatment of covid-19/sars-cov-2 infection: A review describing drug mechanisms of action. *Biochemical pharmacology*, page 114296, 2020.

[110] Keivan Zandi, Franck Amblard, Katie Musall, Jessica Downs-Bowen, Ruby Kleinbard, Adrian Oo, Dongdong Cao, Bo Liang, Olivia O Russell, Tamara McBrayer, et al. Repurposing nucleoside analogs for human coronaviruses. *Antimicrobial Agents and Chemotherapy*, 65(1), 2020.

[111] Qi Zhang, Catherine Zhengzheng Chen, Manju Swaroop, Miao Xu, Lihui Wang, Juhyung Lee, Amy Qiu Wang, Manisha Pradhan, Natalie Hagen, Lu Chen, Min Shen, Zhiji Luo, Xin Xu, Yue Xu, Wenwei Huang, Wei Zheng, and Yihong Ye. Heparan sulfate assists SARS-CoV-2 in cell entry and can be targeted by approved drugs in vitro. *Cell Discovery*, 6(1), November 2020.

[112] Qi Zhang and Sally A Goldman. Em-dd: An improved multiple-instance learning technique. In *Advances in neural information processing systems*, pages 1073–1080, 2002.

[113] Ya-Nan Zhang, Qiu-Yan Zhang, Xiao-Dan Li, Jin Xiong, Shu-Qi Xiao, Zhen Wang, Zhe-Rui Zhang, Cheng-Lin Deng, Xing-Lou Yang, Hong-Ping Wei, et al. Gemcitabine, lycorine and oxysophoridine inhibit novel coronavirus (sars-cov-2) in cell culture. *Emerging microbes & infections*, 9(1):1170–1173, 2020.

[114] Yu Zhao, Fan Yang, Yuqi Fang, Hailing Liu, Niyun Zhou, Jun Zhang, Jiarui Sun, Sen Yang, Bjoern Menze, Xinjuan Fan, et al. Predicting lymph node metastasis using histopathological images based on multiple instance learning with deep graph convolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4837–4846, 2020.

[115] Yadi Zhou, Fei Wang, Jian Tang, Ruth Nussinov, and Feixiong Cheng. Artificial intelligence in covid-19 drug repurposing. *The Lancet Digital Health*, 2020.

[116] Zhi-Hua Zhou, Yu-Yin Sun, and Yu-Feng Li. Multi-instance learning by treating instances as non-iid samples. In *Proceedings of the 26th annual international conference on machine learning*, pages 1249–1256, 2009.

[117] Na Zhu, Wenling Wang, Zhidong Liu, Chaoyang Liang, Wen Wang, Fei Ye, Baoying Huang, Li Zhao, Huijuan Wang, Weimin Zhou, et al. Morphogenesis and cytopathic effect of sars-cov-2 infection in human airway epithelial cells. *Nature communications*, 11(1):1–8, 2020.