

Article

Deep Learning for Drug Discovery: A Study of Identifying High Efficacy Drug Compounds Using a Cascade Transfer Learning Approach

Dylan Zhuang^{1,2} and Ali K. Ibrahim^{1,3,*}¹ EECS Department, Florida Atlantic University, Boca Raton, FL 33431, USA; dzhuang2018@fau.edu² Research Intern from Spanish River High School, Boca Raton, FL 33496, USA³ HBOI Department, Florida Atlantic University, Boca Raton, FL 33431, USA

* Correspondence: aibrahim2014@fau.edu

Abstract: In this research, we applied deep learning to rank the effectiveness of candidate drug compounds in combating viral cells, in particular, SARS-Cov-2 viral cells. For this purpose, two different datasets from Recursion Pharmaceuticals, a siRNA image dataset (RxRx1), which were used to build and calibrate our model for feature extraction, and a SARS-CoV-2 dataset (RxRx19a) was used to train our model for ranking efficacy of candidate drug compounds. The SARS-CoV-2 dataset contained healthy, uninfected control or “mock” cells, as well as “active viral” cells (cells infected with COVID-19), which were the two cell types used to train our deep learning model. In addition, it contains viral cells treated with different drug compounds, which were the cells not used to train but test our model. We devised a new cascade transfer learning strategy to construct our model. We first trained a deep learning model, the DenseNet, with the siRNA set, a dataset with characteristics similar to the SARS-CoV-2 dataset, for feature extraction. We then added additional layers, including a SoftMax layer as an output layer, and retrained the model with active viral cells and mock cells from the SARS-CoV-2 dataset. In the test phase, the SoftMax layer outputs probability (equivalently, efficacy) scores which allows us to rank candidate compounds, and to study the performance of each candidate compound statistically. With this approach, we identified several compounds with high efficacy scores which are promising for the therapeutic treatment of COVID-19. The compounds showing the most promise were GS-441524 and then Remdesivir, which overlapped with these reported in the literature and with these drugs that are approved by FDA, or going through clinical trials and preclinical trials. This study shows the potential of deep learning in its ability to identify promising compounds to aid rapid responses to future pandemic outbreaks.

Keywords: COVID-19; drug discovery; deep learning; cascade transfer learning; DenseNet

Citation: Zhuang, D.; Ibrahim, A.K. Deep Learning for Drug Discovery: A Study of Identifying High Efficacy Drug Compounds Using a Cascade Transfer Learning Approach. *Appl. Sci.* **2021**, *11*, 7772. <https://doi.org/10.3390/app1117772>

Academic Editor: Cecilia Di Ruberto

Received: 6 July 2021

Accepted: 16 August 2021

Published: 24 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In Wuhan, China, following an unknown form of pneumonia outbreak, a new coronavirus emerged, dubbed the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), with the current outbreak being dubbed COVID-19 and promptly declared a “Public Health Emergency of International Concern”. As of the time of writing, there are 200 M coronavirus cases worldwide, with the United States alone being responsible for around 33.6 M cases. The worldwide recorded deaths have surpassed 3.85 M. There are now several different vaccines, and a singular therapeutic drug labeled GS-5734 has passed FDA trials, however its efficacy is contested [1–4].

In the last decades, the cost of drug discovery has increased exponentially, with the median cost of finding an FDA-approved drug at around USD 19 million when writing this article. Due to this, it is prohibitively expensive to evaluate the potential for the treatment of even a proactively curated set of drugs like the one provided by Recursion Pharmaceuticals [5]. Standard preclinical trials involve manual assessment of drug efficacy,

which is hugely time and labor-intensive, while also subject to human error. Due to the time constraints of such a process, the number of drugs that may be analyzed is limited. As such, for a situation such as COVID-19, which both spreads rapidly and has a risk of mutation, any method of streamlining drug discovery can serve to save countless lives. A viable approach in such a case is to use a deep learning model to identify promising compounds and to validate these compounds through clinical trials [6].

Drug discovery using machine learning methods such as support vector machines and neural networks has been proven effective in attempting to lower the prohibitive cost and speed up the process, as opposed to the extensive process of human experimentation [7–11]. In this research, we attempt to demonstrate the potential of deep learning [12,13], a neural network-based machine learning technique, to determine promising drug compounds, using COVID-19 as a proof of concept in order to produce a response for future outbreaks [6,11,12].

Currently, the drug selection process involves target identification and validation, which aims to identify which molecules vary based on the viral status of the cells in order to identify possible gauges for the change caused by drugs [9,14]. Target identification also aims to predict the druggability of a target molecule, or the extent to which the activity of the target compound may be changed through therapeutics. The drug then faces compound screening and lead discovery, which aims to identify promising compounds and drugs through determining which compounds result in the most significant variability of the resultant cells and identifying the most promising compounds as leads, which then proceed to the preclinical and clinical development trials. In this paper, we explore the ability of neural networks to accelerate the first two stages of target identification and validation, as well as compound screening and lead verification. Machine learning has various applications in medicine, with one of the most significant benefits being drug discovery. Biological systems are incredibly complex and contain an immense amount of densely stored information, with the new forms of data allow for even more data for pharmaceutical companies to access and organize. The decision process of the neural network relies on its ability to identify specific characteristics of viral cells and healthy control cells, known as “features,” allowing it to analyze new data and classify the data as “healthy” or “viral,” while relative values of these features associated with each class allow the model to provide a probability score indicating its “confidence” in its classification [5–7,12].

In addition, the broader availability of more powerful computers and deep learning algorithms helps increase the accessibility of deep neural networks (DNNs) [12,13,15,16], and allows for an increasingly robust variety of applications for deep learning within the pharmaceutical industry. There are various effective methods of machine learning, which can be divided into supervised and unsupervised learning techniques. Supervised learning is valuable in its ability to predict labels of data samples, while unsupervised learning is valuable in its ability to categorize and cluster data in ways that the creator may not be aware of. Our model is based on supervised learning, specifically a type of artificial neural network known as a convolutional neural network (CNN). It is well known that different forms of neural networks are suitable in modeling and analyzing different types of data. For example, recurrent neural networks are more capable of analyzing time-variable neural networks, while deep CNNs [3] excel at image processing. Deep learning is useful in its ability to identify biomarkers and build predictive models on the possible effectiveness of a drug, mainly through its ability to identify which drugs lead to the most significant changes.

Despite their apparent value in rapidly accelerating the drug discovery process, it must be noted that neural networks still face challenges in selecting suitable architectures and in data preparation. In order to effectively identify and test targets and compounds, data must be standardized, comprehensive, and of sufficient quantity. Even then, the model is subject to overfitting, where it begins to recognize the noise of the training data, resulting in reduced performance. Overfitting can be combated with techniques such as

regularization, which places a penalty on more complex models, and may also be helped by further processing and cleaning of data [6].

Our focus in this paper is to identify or validate promising compounds and leads for combating viral cells, particularly SARS-CoV-2 viral cells. In the literature, researchers have presented a wide selection of methods for the analysis of Recursion's dataset, with most methods arriving at the same conclusion on the most effective drugs, some of which have also been under clinical trials. For the RxRx19a dataset, there have been several research studies that utilized said dataset. In Recursion's in-house analysis, they concluded that GS-441524 was the clear front runner in terms of preventing cell damage, with Remdesivir, Cx-4945, and Alimitrine all being relatively close to each other in terms of efficacy score [5]. Another study by researchers from the Simon Fraser University and the University of British Columbia found the leading candidates for viral suppression were Remdesivir, GS-441524, and Aloxistatin, as well as the drugs Mebendazole, Oxibendazole, and Albendazole [17]. This group of researchers proposed to use support vector machines to identify potential lead compounds, while our model relies on transfer learning. Our analysis, Recursion's analysis, and analysis from the researchers at Simon Fraser University arrive at similar compounds identified as the most promising, though our approach uses probability as an equivalent efficacy score for each candidate compound.

In this study, we applied transfer learning, a technique that refines a deep learning model pre-trained in one domain for a specific application in another domain, to identify promising compounds and leads for COVID-19. A deep learning model, DenseNet, was first trained with the siRNA image dataset with the transfer learning strategy [18–21]. The trained DenseNet was then refined, retrained, and experimented on the SARS-CoV-2 dataset with 1752 different compounds, also provided by Recursion. In the process, transfer learning is applied twice; therefore, we call the approach "cascade transfer learning." The results obtained by the deep learning model were then compared with those obtained by other models and with those approved by FDA and under clinical trials. The experimental study shows that the cascade transfer learning approach is promising in identifying promising compounds and leads. The remainder of the paper is organized in the following manner. Section 2 introduces the SARS-CoV-2 dataset from Recursion. Section 3 details the deep learning strategy adopted in this research. Section 4 presents the experimental results with discussions. The paper ends with concluding remarks.

2. Dataset

In order to work as efficiently as possible, we utilized a morphological imaging dataset from Recursion Pharmaceuticals with 305,520, 5-channel images. In April of 2020, Recursion developed the dataset, where cells were treated in 1536 various microplates, with fixation through 5% paraformaldehyde, permeabilization through 0.25% Triton, and staining with a variety of different dyes. Three different cell treatments were involved, with mock cells, which served as a control, cells with irradiated and, therefore, deactivated SARS-CoV-2 virions, and cells infected with active SARS-CoV-2 at a ratio of 4 virions to 10 cells. These cells were then treated with a large number of candidate compounds, which came from a library of drugs including FDA approved drugs, European Medicines Agency (EMA) approved drugs, and compounds in clinical trials for SARS, which were compiled into the RxRx19a dataset [17]. Each compound was given in concentrations, varying in 6 different concentrations, each separated by half-log increments, with each concentration of each compound being repeated 6 times. The SARS-CoV-2 virus targets human kidney cells, with EM observations indicating the presence of virus-like particles within the kidney cells of coronavirus infected patients. The cytopathic nature of the SARS-CoV-2 virus suggests direct tubular damage through cytotoxicity. This knowledge is applied through the usage of continuous renal replacement therapies on COVID-19 patients due to the observation of acute kidney injury in patients, but was also crucial in the creation of the RxRx19a dataset, as tubule damage meant that the infected kidney cells of a COVID-19 patient possess characteristics that would distinguish them from mock cells. Recursion used human renal

cortical epithelial cells, or HRCE cells from proximal and distal tubules, as well as VERO cells from the kidney cells of the African green monkey as a control in the dataset. An example of the images described above can be found in Figure 1 [5].

Recursion also released another set of data with the same treatment conditions named RxRx1, but with a 384-well plate density, as opposed to the 1536 plate density of RxRx19, an image resolution of 512×512 with 6 channels instead of the image resolution of 1024×1024 with 5 channels of RxRx19, and 1138 classes of siRNA as opposed to the 1672 compounds of RxRx19. The cells used in the previous experiment were treated with siRNA instead of target compounds. The siRNA dataset shared similar features to the SARS-CoV-2 dataset, but with 1139 classes of siRNA applications instead of the 1672 molecules and three viral conditions of the RxRx19a dataset. This allows us to use the siRNA dataset, a relatively larger dataset, to train the classification model first, and then use the viral/mock cells in the SARS-CoV-2 dataset to retrain the model [5]. Details on this will be given in Sections 3.3 and 4.

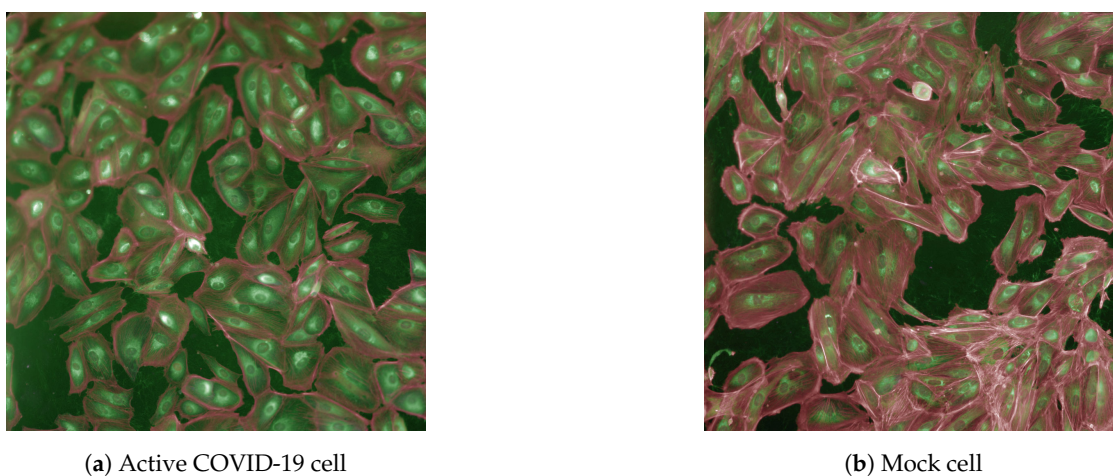


Figure 1. An active-viral cell image and a mock cell image.

3. Transfer Learning

Most classification tasks can be divided into feature extraction and classification stages, where feature extraction determines the meaningful and non-redundant features of a dataset. In this study, we used the strategy of transfer learning for feature extraction.

3.1. Introduction to Transfer Learning

Transfer learning is a method where a model and its weights developed for one task is used as a primary building block for another related task [18]. An illustration of transfer learning is provided in Figure 2. Transfer learning is notably valuable for image identification and other related tasks due to the training time required [19–21]. The core steps for transfer learning are taking layers from another model, freezing said layers, adding classification layers, and training the new layers on the specified dataset. There is also an optional step for training after the classification layers, where the weights for the transferred model are unfrozen and retrained on the dataset for an additional improvement in performance. This process is known as fine-tuning [2].

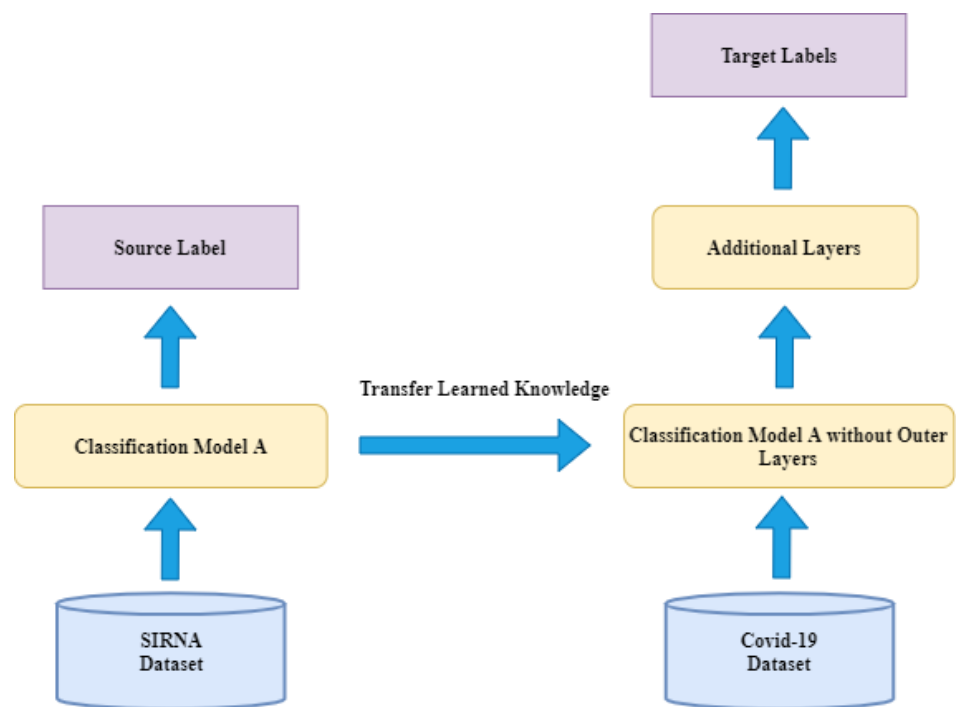


Figure 2. Transfer learning concept.

3.2. Convolutional Neural Networks

CNNs, proven to be some of the best tools in image analysis [22], are often used in transfer learning. An activation function like ReLu is used to take the complex values from the network and simplify it into an output. An illustration of CNN is provided in Figure 3.

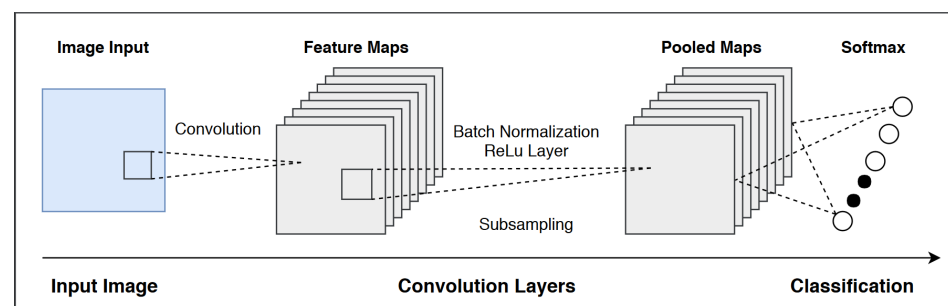


Figure 3. General architecture of a convolutional neural network.

Inputs to CNNs are first fed into the convolution layers, which uses groups of adjustable filters, with the most common parameters being convolution kernel size, step, and filter quantity in order to detect specific patterns inside of an input layer [3]. Activation functions, as well as batch normalizations before or after the activation functions, are applied in order to simplify the outputs from the convolution, and finally, a pooling and fully connected or SoftMax layers are applied in order to synthesize the outputs of the convolution layers. In CNNs, the convolution kernels and filters can be adjusted along with the weights and biases through backpropagation and gradient descent [3].

In this study, we used a pre-trained CNN, the DenseNet-161 architecture [23]. It connected every layer to every other layer to capitalize on the fact that models perform better at the beginning and end of the model. For each layer in the DenseNet architecture, the feature maps of previous layers were used as inputs for the layers that follow them. DenseNets reduces problems such as the vanishing gradient issue, improves feature propagation, and encourages feature reuse while requiring much less computational power with relatively high performance.

The DenseNet 161 architecture we adopted for the study was trained with the 1139-class dataset of siRNA images. It featured an input layer, with an input and output size of $512 \times 512 \times 6$ [23]. It was followed by a convolution layer, which used a convolution filter, a max pool layer, a dense block, and a transition layer. This structure was repeated four times and then followed by a pooling layer and fully connected layer with SoftMax primarily for feature extraction. The fully-connected and batch normalization layers were used to standardize the layer inputs and coordinate layer updates to streamline the training process.

3.3. Cascade Transfer Learning

We used the cascade transfer learning strategy to classify the five input channels from the RxRx19a dataset to mock cells and active SARS-CoV-2 viral cells, as shown in Figure 4. First, the pre-trained DenseNet was trained with the siRNA images, as shown at the top of the figure. From the figure, it can be seen that the DenseNet acted as a feature extractor. Next, the last three layers of the trained DenseNet, a batch normalization (BN) layer, a fully connected (FC) layer, and a SoftMax layer (shown in the top-right of the figure), were replaced by a FC layer, a BN layer, another FC layer, another BN layer, a third FC layer, and then finally a SoftMax layer (shown in the bottom-right of the figure). The number of neurons in the added layers was mainly constrained by the input and output sizes. It must be noted that the size of the SARS-CoV-2 images in the RxRx19a dataset was 1024×1024 , therefore, a down-sampling procedure was applied to these images so they can be processed by the feature extraction model, i.e., the trained DenseNet in the first stage. The scoring model shown at the bottom of the figure was then retrained with the RxRx19a dataset. After the layers in the scoring model were trained with the COVID-19 dataset, we fine-tuned the entire classification model by unfreezing the weights of the DenseNet. The Adam Optimizer, an optimization algorithm for stochastic gradient descent provided by MATLAB, was used to train the deep learning model in both stages. Additional illustration of the idea is given in Figure 4.

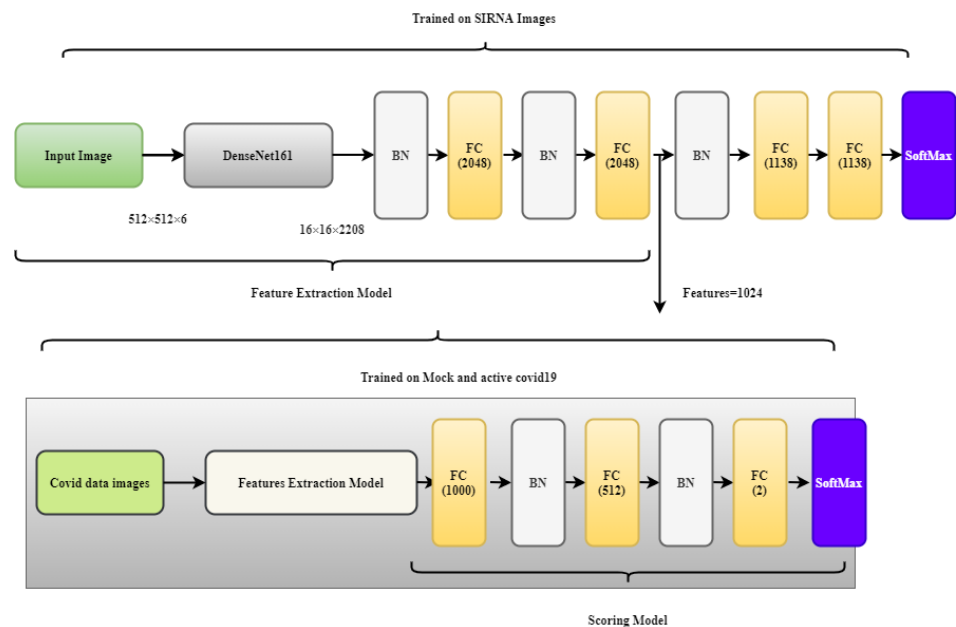


Figure 4. Architecture of the COVID-19 drug development model.

The data from the RxRx1 dataset shared many characteristics with the RxRx19a dataset, which was suitable for transfer learning. As such, we pre-trained the DenseNet model for feature extraction and dropped the classification layers, instead opting for classification layers that would classify the cells into “healthy” and “viral/active,” with a normalized exponential function to provide an efficacy score for each of the candidate compounds.

A healthy cell is assigned a score of “0”, and a viral cell with no treatment is assigned a score of “1”. Therefore, mock cells are labeled “0,” and active viral cells are labeled “1”. A viral cell treated by a compound is scored between 0 and 1, indicating the degree of effectiveness of the compound in treating COVID-19. Thus, with our method, we could take the mean of the outputs for each compound, which could give us a more informative measure for the perceived impact of each compound instead of a binary “healthy/active” output. This method allows us to select the compounds with the most promising leads for further investigation.

4. Experimental Study

The experiments were implemented using the MATLAB software on a computer server Nvidia DGX Workstation with 256 GB of RAM. First, we created a model to classify siRNA images based on the DenseNet161, and we then compared this model with different pre-trained models, such as VGG10, AlexNet, and GoogleNet. Table 1 shows the classification results of different models using siRNA image datasets. It is clear from the table that the DenseNet produced the best accuracy performance.

Table 1. Experimental results with different pre-trained models on siRNA dataset.

Deep Learning Model	Accuracy
VGG16	81.2%
GoogleNet	83.1%
VGG19	83.3%
AlexNet	77.4%
DenseNet	96.4%

We applied a 5-fold validation procedure to check the performance of the model, as shown in Figure 5. In this procedure, 80% of the data were randomly chosen for training and the remaining 20% for validation. This step was repeated five times until all mock/active viral cell images in the dataset were validated once. The results from these five tests were averaged. It must be stressed that in the training dataset, there were only active viral cells and mock cells. There were viral cells treated with different compounds in the test dataset, such as Remdesivir and GS-44152. We compare the effectiveness of our model with other well-known pre-trained models in classifying mock and active viral cells. Table 2 shows the results produced by these models in classifying mock/active viral cells.

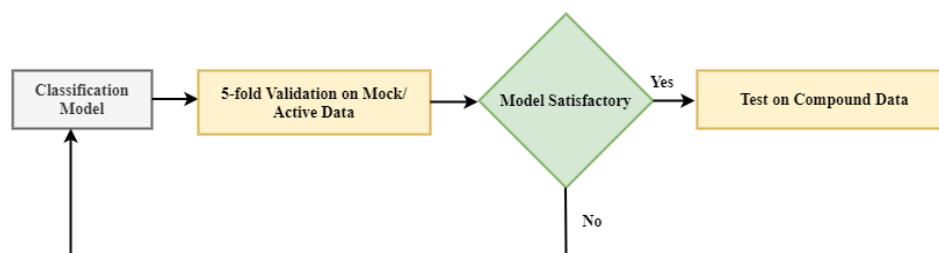


Figure 5. Simplified model test procedure based on the USFDA MAQC II protocol [24].

Sensitivity serves as a benchmark for the accuracy of a test’s positive prediction, while specificity measures the accuracy of a test’s null prediction [25]. In the context of our study, sensitivity is the ability of a particular model to determine active viral cells, while specificity is its ability to determine mock/control cells. The F1 score statistically measures the balance of sensitivity and specificity, indicating the ability of the model in preventing both false positives and false negatives [26]. Finally, the Kappa score is a way of statistically measuring the agreement throughout all the tested cases on a scale of less than and equal to 1, with less than 0 implying little agreement while 1 indicates perfect agreement [27]. Referencing Table 2, the cascade model outperformed other popular deep

neural network architectures. In comparison to the two popular deep learning models for the RxRx19 dataset, our cascade transfer learning model had the highest performance in terms of the given metrics. It must be noted that the vgg19 and GoogleNet models were pre-trained on the ImageNet dataset instead of the siRNA dataset before being retrained on the SARS-CoV-2 dataset. To further study the model's performance, its receiver operating characteristic (ROC) curve is plotted (refer to Figure 6). The area under the curve (AUC) is 0.98.

Table 2. Experimental results with different pre-trained models on SARS-CoV-2 dataset.

Deep Learning Model	Sensitivity	Specificity	F1-Score	Kappa
vgg19	0.86	0.98	0.90	0.74
GoogleNet	0.84	0.97	0.84	0.70
Cascade Transfer Learning	0.97	0.99	0.97	0.87

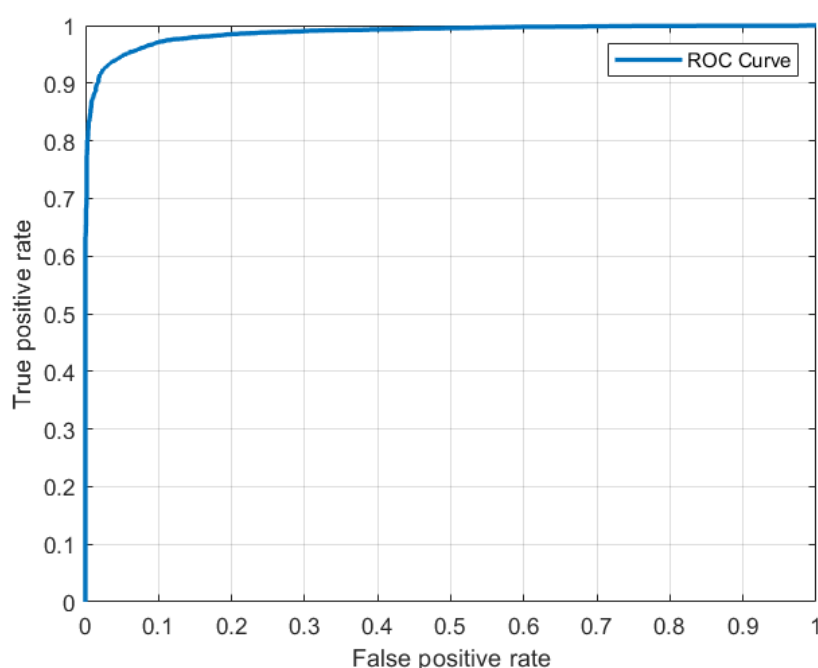


Figure 6. ROC curve of the cascade transfer learning model. The AUC is 0.98.

We then tested the cascade transfer learning model by using it to rank the efficacy of the compounds used to treat COVID-19, as shown in the last block of Figure 5, with a score less than 0.5 indicating the promise of the drug as a lead. These test results are given in Tables 3 and 4 and Figure 7. In the figure, a low probability value, or equivalently a high efficacy score obtained by 1-probability, indicates that the corresponding compound effectively combats active viral cells. In each plot, the histograms of one high efficacy compound and one low efficacy compound are provided. The drugs identified to have the greatest impact on cell health are GS-441524, Remdesivir, CX-4945, Aloxistatin, and Calcipotriene. Of these drugs, two are anti-viral drugs, specifically GS-441524 and Remdesivir, with GS-441524 being a primary component of Remdesivir, explaining their similarly low probability (or high efficacy) scores. In terms of the histograms shown in Figure 7, GS-441524 provides more consistent performance. The next two lowest probability scores were CX-4945 and Aloxistatin, which were both protein inhibitors. The drugs with the highest probability of classifying as Active SARS-Cov-2 were corticosteroids and regulatory drugs, which had little relation to the Cytopathic nature of COVID-19. The drugs identified in this research are similar to the drugs identified by Recursion in their study [5] and by the study reported in [17], as all resulted in GS-441524 and Remdesivir being isolated as clear outliers in our probability scores and also their efficacy scores. Notably, our CX-4945 probability

score is much more noticeable for the same compound, although all results share the same four most effective drugs, albeit in different rankings and orders of magnitude [28]. Of the identified compounds, CX-4945 and Aloxistatin have entered the observational clinical pretrial phase, while Remdesivir (and by association GS-441524) has been approved by the FDA as the first treatment for COVID-19.

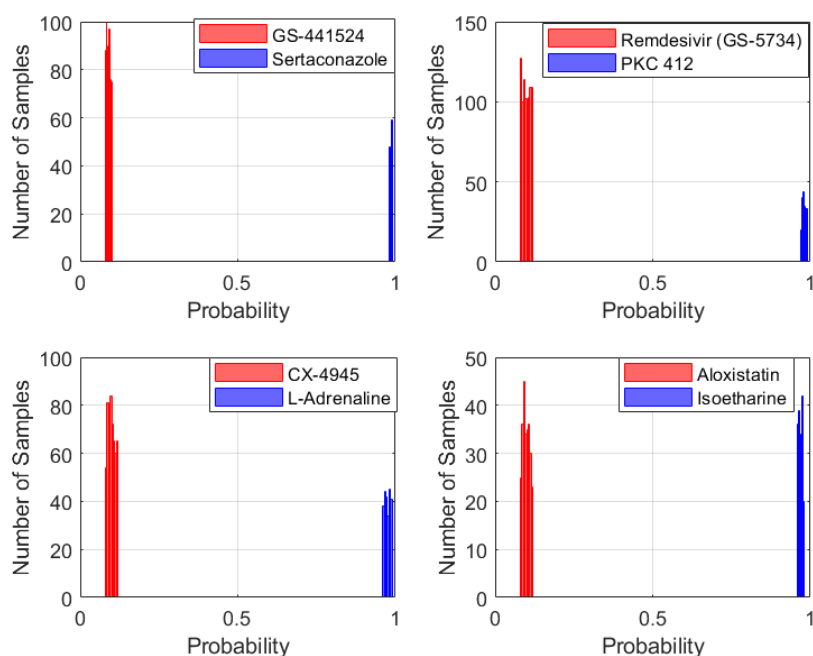


Figure 7. Histograms of probability scores for different compounds. In each plot, for each compound, the horizontal axis is divided in to 7 bins. The vertical axis shows the number of samples in each bin.

Table 3. List of compounds with low probability scores (classified as mock).

Compound	Probability
GS-441524	0.08
Remdesivir (GS-5734)	0.08
CX-4945	0.13
Aloxistatin	0.18
Calcipotriene	0.21

Table 4. List of Compounds with High Probability Scores (Classified as Active Virus).

Compound	Probability
Sertaconazole	0.98
PKC 412	0.97
L-Adrenaline	0.97
Isoetharine	0.96
desoximetasone	0.95

5. Conclusions

This paper presents the preliminary results of an exploratory research on the feasibility of deep learning to predict the efficacy of compounds for drug discovery, using COVID-19 as a case study. A transfer learning approach was adopted for this task.

One novelty of the proposed approach is the way the transfer learning strategy is implemented. We first trained the DenseNet, a pre-trained deep neural network, with the siRNA dataset, larger than the SARS-CoV-2 dataset with similar characteristics. The resulting model is then used to extract features from mock cells and active viral cells provided in

the SARS-CoV-2 dataset. Thus, we used transfer learning twice from the ImageNet to the siRNA image dataset and then from the siRNA dataset to the SARS-CoV-2 dataset. This cascade transfer learning approach produced superior results for the case study. Another novelty of the approach is using a SoftMax layer as the output layer for the classifier, which produces probability (equivalently efficacy) scores for classifying viral cells treated with different compounds, which allows users to analyze test results with a statistical method. Experimental results demonstrated that the model was able to identify highly promising compounds, which were consistent with those identified in the literature and the drugs that are now under clinical trials.

It must be noted that though the proposed transfer learning strategy is more explainable compared with a classification approach that outputs a hard binary decision, owing to the fact that it produces efficacy scores for candidate compounds in treating viral cells, the model on the whole is not yet transparent. For example, we have not yet been able to explain the properties of the features linked to either viral cells or mock cells. A way to improve the model's explainability is to link the features extracted by the model to biomarkers of mock and viral cells with a functional relationship which may also be realized by a data-driven model.

Although this paper is limited in its applicability due to the requirements of obtaining a dataset to train the deep learning model, this research encourages us. The deep learning model trained only on the infected and healthy cells was then able to rank candidate compounds, among which some promising ones were under clinical trials. In theory, transfer learning serves to accelerate a drug discovery process significantly. It produces a list of compounds relatively effective on viral cells, thus providing a baseline for speeding up preclinical trials. With cases growing exponentially in a pandemic, the longer treatments and preventative measures develop, the more the situation worsens. Deep learning is a viable tool to combat future COVID-19-like outbreaks.

Author Contributions: Conceptualization, A.K.I. and D.Z.; methodology, A.K.I. and D.Z.; software, A.K.I. and D.Z.; validation, A.K.I.; formal analysis, D.Z. and A.K.I.; writing—original draft preparation, D.Z.; writing—review and editing, D.Z. and A.K.I.; visualization, A.K.I. and D.Z.; supervision, A.K.I.; project administration, A.K.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this research was provided by Recursion Pharmaceuticals.

Acknowledgments: The authors acknowledge Recursion Pharmaceuticals for providing the RxRx19a dataset for public usage, as well as the RxRx1 dataset, which we used to calibrate our initial feature extraction model. The authors also wish to express their gratitude to the anonymous reviewers and Chuck Cooper of Boca Raton, Florida for their constructive comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Geleris, J.; Sun, Y.; Platt, J.; Zucker, J.; Baldwin, M.; Hripcsak, G.; Labella, A.; Manson, D.K.; Kubin, C.; Barr, R.G.; et al. Observational study of hydroxychloroquine in hospitalized patients with Covid-19. *N. Engl. J. Med.* **2020**, *382*, 2411–2418. [[CrossRef](#)] [[PubMed](#)]
2. Torrey, L.; Shavlik, J. *Transfer Learning*; University of Wisconsin: Madison, WI, USA, 2010; pp. 242–264.
3. Kalchbrenner, N.; Grefenstette, E.; Blunsom, P. A convolutional neural network for modelling sentences. *arXiv* **2014**, arXiv:1404.2188.
4. Chang, L.; Yan, Y.; Wang, L. Coronavirus disease 2019: Coronaviruses and blood safety. *Transfus. Med. Rev.* **2020**, *34*, 75–80. [[CrossRef](#)] [[PubMed](#)]

5. Heiser, K.; McLean, P.F.; Davis, C.T.; Fogelson, B.; Gordon, H.B.; Jacobson, P.; Hurst, B.L.; Miller, B.J.; Alfa, R.W.; Earnshaw, B.A.; 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**. [CrossRef]
6. Kandoi, G.; Acencio, M.L.; Lemke, N. Prediction of druggable proteins using machine learning and systems biology: A mini-review. *Front. Physiol.* **2015**, *6*, 366. [CrossRef] [PubMed]
7. Urban, G.; Bache, K.; Phan, D.T.; Sobrino, A.; Shmakov, A.K.; Hachey, S.J.; Hughes, C.C.; Baldi, P. Deep learning for drug discovery and cancer research: Automated analysis of vascularization images. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2018**, *16*, 1029–1035. [CrossRef] [PubMed]
8. Akondi, V.S.; Menon, V.; Baudry, J.; Whittle, J. Novel K-Means Clustering-Based Undersampling and Feature Selection for Drug Discovery Applications. In Proceedings of the 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 18–21 November 2019; pp. 2771–2778.
9. Farag, A.; Wang, P.; Ahmed, M.; Sadek, H. Identification of FDA Approved Drugs Targeting COVID-19 Virus by Structure-Based Drug Repositioning. 2020. Available online: <https://chemrxiv.org/engage/chemrxiv/article-details/60c74b2a567dfe0f38ec4ee7> (accessed on 1 August 2021).
10. Bray, M.A.; Singh, S.; Han, H.; Davis, C.T.; Borgeson, B.; Hartland, C.; Kost-Alimova, M.; Gustafsdottir, S.M.; Gibson, C.C.; Carpenter, A.E. Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat. Protoc.* **2016**, *11*, 1757. [CrossRef]
11. Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M.; et al. Applications of machine learning in drug discovery and development. *Nat. Rev. Drug Discov.* **2019**, *18*, 463–477. [CrossRef]
12. Chen, H.; Engkvist, O.; Wang, Y.; Olivecrona, M.; Blaschke, T. The rise of deep learning in drug discovery. *Drug Discov. Today* **2018**, *23*, 1241–1250. [CrossRef]
13. Hinton, G. Deep learning—A technology with the potential to transform health care. *JAMA* **2018**, *320*, 1101–1102. [CrossRef]
14. Jeon, J.; Nim, S.; Teyra, J.; Datti, A.; Wrana, J.L.; Sidhu, S.S.; Moffat, J.; Kim, P.M. A systematic approach to identify novel cancer drug targets using machine learning, inhibitor design and high-throughput screening. *Genome Med.* **2014**, *6*, 1–18. [CrossRef] [PubMed]
15. Le Cun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
16. Wang, J.; Wei, Z.; Zhang, T.; Zeng, W. Deeply-fused nets. *arXiv* **2016**, arXiv:1605.07716.
17. Saberian, M.S.; Moriarty, K.P.; Olmstead, A.D.; Nabi, I.R.; Jean, F.; Libbrecht, M.W.; Hamarneh, G. DEEMD: Drug Efficacy Estimation against SARS-CoV-2 based on cell Morphology with Deep multiple instance learning. *arXiv* **2021**, arXiv:2105.05758.
18. Driessens, K.; Ramon, J.; Croonenborghs, T. Transfer Learning for Reinforcement Learning through Goal and Policy Parametrization. 2006; pp. 1–4. Available online: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.490.9085&rep=rep1&type=pdf> (accessed on 1 August 2021).
19. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345–1359. [CrossRef]
20. Bengio, Y. Deep learning of representations for unsupervised and transfer learning. In *Proceedings of the ICML Workshop on Unsupervised and Transfer Learning, JMLR Workshop and Conference Proceedings*; PMLR: Bellevue, WA, USA, 2012; pp. 17–36.
21. Ibrahim, A.K.; Zhuang, H.; Chérubin, L.M.; Schärer-Umpierre, M.T.; Nemeth, R.S.; Erdol, N.; Ali, A.M. Transfer learning for efficient classification of grouper sound. *J. Acoust. Soc. Am.* **2020**, *148*, EL260. [CrossRef]
22. Godinez, W.J.; Hossain, I.; Lazić, S.E.; Davies, J.W.; Zhang, X. A multi-scale convolutional neural network for phenotyping high-content cellular images. *Bioinformatics* **2017**, *33*, 2010–2019. [CrossRef]
23. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
24. Shi, L.; Campbell, G.; Jones, W.D.; Campagne, F.; Wen, Z.; Walker, S.J.; Su, Z.; Chu, T.M.; Goodsaid, F.M.; Pusztai, L.; et al. The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. *Nat. Biotechnol.* **2010**, *28*, 827–838.
25. Baratloo, A.; Hosseini, M.; Negida, A.; El Ashal, G. Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity. *Emergency* **2015**, *3*, 48–49.
26. Sasaki, Y. The Truth of the F-Measure. 2007. Available online: <https://www.toyota-ti.ac.jp/Lab/Denshi/COIN/people/yutaka.sasaki/F-measure-YS-26Oct07.pdf> (accessed on 1 August 2021).
27. McHugh, M.L. Interrater reliability: The kappa statistic. *Biochem. Med.* **2012**, *22*, 276–282. [CrossRef]
28. Khalifa, N.E.M.; Taha, M.H.N.; Manogaran, G.; Loey, M. A deep learning model and machine learning methods for the classification of potential coronavirus treatments on a single human cell. *J. Nanoparticle Res.* **2020**, *22*, 1–13. [CrossRef] [PubMed]