COMP 4471 Final Report: Detecting COVID-19 via Chest X-rays Using Deep Learning

SUN Yushi
The Hong Kong University of Science and Technology
The Clear Water Bay, Hong Kong
ysunbp@connect.ust.hk

LI Tianle
The Hong Kong University of Science and Technology
The Clear Water Bay, Hong Kong
tliax@connect.ust.hk

Abstract

The novel corona-virus disease(COVID-19) is an ongoing pandemic, which has infected more than 4.5 million people and caused fatalities for around 300 thousand[27]. The goal of our project is to develop a deep learning-based model to detect the COVID-19 patients through their Chest X-ray scans. Our project utilized the transfer learning and model ensemble techniques in deep learning to train the existing CNN models on the datasets posted on Kaggle[9][11][22]. Our model achieved over 97% accuracy with high precision, sensitivity(recall) and specificity in the task of classifying pneumonia X-ray scans from the healthy X-ray scans on the adjusted dataset provided by the University of California San Diego[22] and over 98% accuracy with high precision, sensitivity(recall) and specificity in the task of classifying COVID-19 X-ray scans from the pneumonia X-ray scans on the combined dataset provided by QatarUniversity, the University of Dhaka[9] and the University of Montreal[11], which are believed to out-perform the current state-of-art models and are useful for the reference of real-life diagnosis. Apart from the classification models, we also developed an automated bone-suppressed, lobe-enhanced model based on CycleGAN[5][10], which can assist with the manual diagnosis and other classification tasks for this pandemic[16].

1. Introduction

According to a study about COVID-19, this disease presents several unique imaging features[8]. The Chest X-ray scans of the infected patients may present some useful characteristic features, which are hardly visible to human eyes[26]. In order to accelerate the diagnosis of this rapidly spreading disease. We need to propose an automated COVID-19 Chest X-ray detecting CNN model to help doctors improve their diagnosis speed and accuracy. The reason why our work is important is that X-ray scans is more accessible than other diagnosis methods such as CT scans and RT-PCR testing in rural area, thus a model that can classify COVID-19 X-ray scans from healthy X-ray scans would be of great help to these people. Besides, the detour of our major work: the bone-suppress model can be of great help for doctors to spot the nidi and nodules from the X-ray scans, thus improve the accuracy of pathologies classification[16].

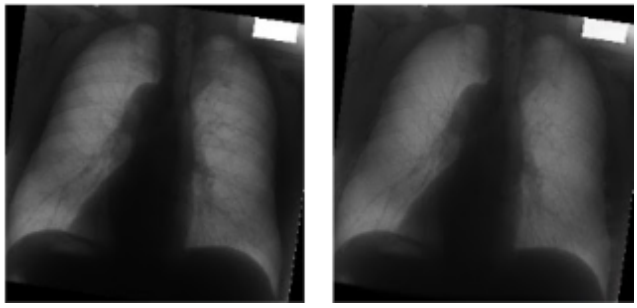


Figure 1: Demonstration of the result generated by our bone-suppress model: The original Chest X-ray scan (left) and the bone-suppressed version of that Chest X-ray scan (right).

A brief overview of our result is as following: In the first phase of our model, which is to classify pneumonia X-ray scans from the healthy X-ray scans, our model achieved

97.87% accuracy with 97.67% precision, 99.41% sensitivity(recall) and 94.80% specificity; In the second phase of our model, which is to classify COVID-19 X-ray scans from the pneumonia X-ray scans, our model achieved 98.59% accuracy with 96.15% precision, 96.15% sensitivity(recall) and 99.13% specificity. A more detailed illustration would be shown in the **5. Experiment** section. For the detour of our model, as shown in Figure 1, the shadow cast by ribs has been significantly suppressed while the lung lobe tissue has been well preserved.

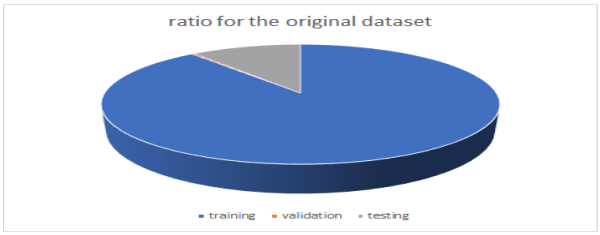
2. Related Work

2.1. Existing Medical Detection Methods

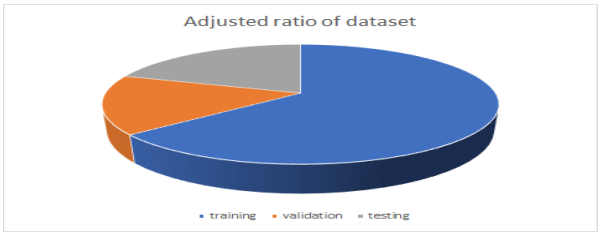
As a gold standard in detecting repository specimens[4], Reverse Transcription Polymerase Chain Reaction (RT-PCR) is currently adopted as the main diagnosis method[30], however, this method has its potential drawbacks: time-consuming, shortage of supply, complicated process with a relatively low positive rate of around 63%[30]. Another main diagnosis method used is the radiography-based clinical symptoms diagnosis[31]. Patients are diagnosed through their Computed Tomography scans(CT) or Chest X-ray scans. However, this method suffers from the difficulty that the radiographic differences between COVID-19 scans and other viral pneumonia and bacterial pneumonia scans are subtle and hard to capture by naked eyes[9]. According to a study of differentiating COVID-19 scans from viral pneumonia scans, most radiologists present an accuracy lower than 90% in distinguishing COVID-19 CT scans[6]. Thus, the radiography-based clinical symptoms diagnosis method alone may not achieve a perfect performance in solving this problem.

2.2. Deep Learning Based Methods

In order to assist with the existing clinical diagnosis methods, several deep learning based diagnosis methods have been proposed recently: Wang *et al.*[29] proposed COVID-NET designed for detecting COVID-19 based on Chest X-ray scans. The key difference between COVID-NET and our ensembled model is that COVID-NET utilizes residual architecture and our ensembled model focuses on balance the performance between VGG-16 model and SqueezeNet model to achieve a better classification accuracy. Another difference is that COVID-NET classifies the X-ray scans into healthy, non-COVID and COVID-19 classes, while our model conquers the same classification task by dividing the whole classification task into two phases. With this 'divide-and-conquer' spirit, our model achieves a higher overall accuracy. Chowdhury *et al.*[9] also examined the possibilities to apply deep learning CNN models on classifying COVID-19 based on Chest X-ray scans. They utilized the pre-trained DenseNet201[18], Resnet18[17], AlexNet[24] and SqueezeNet[20]. The similar points between our model and this work is that both of us evaluated the performance for Resnet and SqueezeNet. A key difference is that Chowdhury *et al.* only focused on tuning single model but not ensembling different models to achieve a better performance. Satisfactory results have been demonstrated by other works proposed by Gozes *et al.*[15], Li *et al.*[25] and Butt *et al.*[7] applied deep learning methods on the CT scans. The difference between our model and theirs is obvious: they trained their model on CT scans while we trained our model on Chest X-ray scans.



(a) Original distribution



(b) Adjusted distribution

Figure 2: Redistribution of dataset

3. Data

3.1. Data Overview

We obtained our Chest X-ray scans dataset from the works by Kermany *et al.*[22], Chowdhurry *et al.*[9] and Cohen *et al.*[11], a demonstration of the images in the dataset is shown in Figure 3. Due to the different tasks we defined in Phase 1 and Phase 2, we used different datasets for these two phases.
The dataset used for our detour: the bone-suppress model is attained from[19], which contains 8642 Chest X-ray scans images with half of them unchanged and half of them bone-suppressed.

3.1.1 Phase 1

In phase 1 our task is to classify pneumonia X-ray scans from healthy X-ray scans, thus we opt for the dataset provided by Kermany *et al.*[22], which contains 5856 images with 4273 healthy scans and 1583 pneumonia scans. A notable point is that the original distribution of training, val-

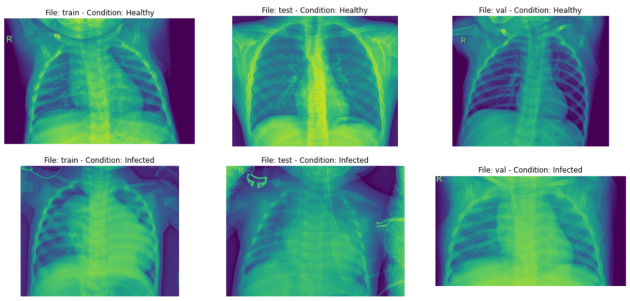


Figure 3: A demonstration of the X-ray scans we used in our models.

idation and testing images are quite unbalanced: 5216 images in training set, 16 in validation and 624 in testing set (as shown in Figure 2a). So an adjustment step is performed in **3.2 Data Pre-processing** section to achieve a better partition of the dataset for our training and evaluation.

3.1.2 Phase 2

In phase 2 our task is to distinguish COVID-19 scans from the pneumonia scans (both bacterial and viral types). We then combined the three datasets[9][11][22] to create a new dataset for our task. The new dataset we created contains 2275 images in total with 219 COVID-19 scans and 1345 viral pneumonia scans provided by Chowdhury *et al.*[9], 125 COVID-19 scans from Cohen *et al.*[11] and 586 bacterial scans from Kermany *et al.*[22]. We partition the images into training, validation and testing set with a ratio of 8:1:1, COVID-19 scans and Peumonia scans are evenly distributed among the three sets.

3.1.3 Detour: Bone-Suppression

In the bone suppression task, we obtained the original X-ray scans and their bone-suppressed target from Kaggle[19]. We separated the dataset into training set and testing set with a small proportion assigned to testing, as we are able to tell the effect by naked eyes.

3.2. Data Pre-processing

3.2.1 Phase 1 & 2

In this stage, all images are resized to 100×100 pixels in phase 1 and 150×150 pixels in phase 2 as we found it merely triggered trivial difference for test results with different resize scale within all models we chose. Thus, we normalized the X-ray scans in the same way to make it more convenient for us to implement model ensemble in the final stage.

Besides, we redistributed the original dataset[22] for phase 1 by changing the ratio of training, validation and testing sets, because there are only 16 images in the validation set, only coming from normal class and bacteria pneumonia class but not virus pneumonia class. After the adjustment, the proportions of training, validation and testing sets become 64%, 16% and 20% respectively with all classes evenly distributed as shown in Figure 2.

3.2.2 Detour: Bone-Suppression

The pre-augmented version of the X-ray scans with random rotation and flipping are provided by[19]. Thus, we directly exploited the augmented ones and scaled the images from [0, 225] to [-1, 1] typically for GAN generator models before feeding them into our CycleGAN model.

4. Methods

We decided to proceed our project gradually, thus divided the whole project into two phases. In Phase 1, we only focus on classifying the pneumonia scans from normal scans. In Phase 2, we step further to classify COVID-19 scans from the other virus pneumonia scans. Then, we added bacterial pneumonia scans from the dataset we used in phase1[22] and heterogeneous COVID-19 scans from another source [11] to retrain our model so as to boost the model robustness towards multitudinous inputs format. Moreover, we transferred the original X-ray scans to a bone-suppressed version that can improve the visual quality of the original scans and make the pneumonia symptoms feature clearer to naked eye.

4.1. Methodology in Phase1 and Phase2

Generally, CNN requires a large number of training datasets to outperform other machine learning algorithm and it usually takes quite a long time to train the deep learning model until it converges to the optimal stage. In the condition of limited time and relatively small scale published chest X-ray dataset[9][22], we decided to choose transfer learning as our first resort and utilize the weights of the models trained on ImageNet[12] initially. Ultimately, we ensemble some of the well-trained models to achieve a better comprehensive test results.

In this study, we exploited three well-known pre-trained deep learning CNNs: VGG-16[28], ResNet-50[17] and SqueezeNet[20] to classify the chest x-ray images for two classification problems in phase1 and phase2. The first two of them was firstly weighted using transfer learning from ImageNet[12] by Keras API, while we trained the SqueezeNet model from scratch.

The VGG-16 architecture contains twelve convolutional layers, some of which are followed by maximum pooling layers and then four fully-connected layers and finally a 1000-way softmax classifier. The detailed architecture can be shown in Figure 4.

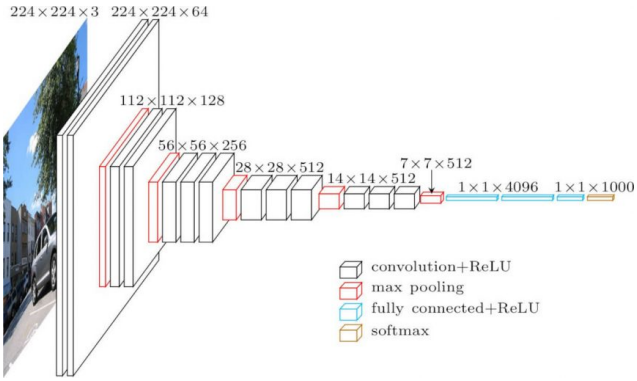


Figure 4: VGG16 Architecture

The residual network[17] is based on a plain network similar with VGG net, and a shortcut connection is inserted which turns the network into its counterpart residual version. Especially for ResNet-50[17], it consists of 5 stages each with a convolution and Identity block. Each convolution block has 3 convolution layers and each identity block also has 3 convolution layers as shown in Figure 5. Although it is deeper than VGG-16, the parameter size is much smaller than VGG nets.

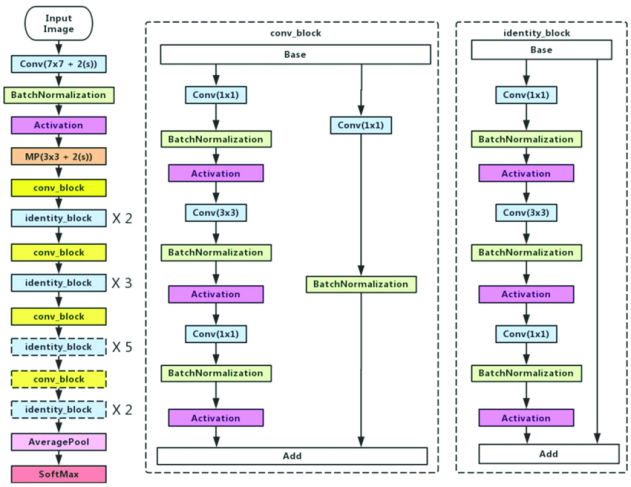
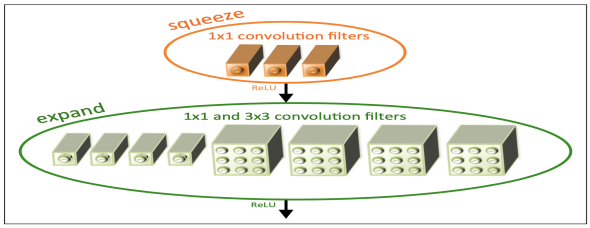
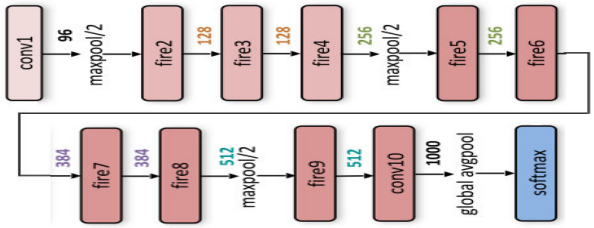


Figure 5: ResNet-50 Architecture

Moreover, we also utilized SqueezeNet[20] as a relatively efficient model to train our data. It begins with a single convolution layer (conv1), followed by 8 fire modules (fire2–9), and ends with a final convolution layer (conv10), where the fire module consists of two layers: a squeeze layer and an expand layer. Both of them will keep the same feature map size, while the former reduces the depth to a smaller number, and the later increases it. We can adjust the number of filters per fire module accordingly to a specific task. The detailed architecture can be found in Figure 6.



(a) Fire module architecture



(b) SqueezeNet Architecture

Figure 6: Overall SqueezeNet Architecture

Finally, we implemented model ensemble[13] with the best weighted VGG-16 and SqueezeNet we trained to achieve a higher test accuracy score. The reason why we left out ResNet-50 will be explained in next section.

4.2. Methodology in Bone Suppression

In this part, we chose CycleGAN[10] as our model, because as an image-to-image translation task, most of the cases the paired training dataset is generally not available, especially in the medical field. And CycleGAN can solve this problem with a cycle consistence loss mentioned in[10].

A CycleGAN consists of two generators and two discriminators as shown in Figure 7. The discriminators D_X and D_Y classify an input chest X-ray scan as real or fake. D_Y promotes the generator $G_{X\to Y}$ to learn the mapping $X \to Y$ and translate original X-ray scans from domain X into outputs indistinguishable from target domain Y (bone-suppressed version in our case) , and vice versa for D_X and $G_{Y\to X}$. Beyond adversarial losses[14], forward and backward cycle-consistency loss[32] are also utilized

Task	Model	Accuracy	Sensitivity(Recall)	Precision	Specificity
Normal and Pneumonia before data redistribution	VGG-16	0.9199	0.9846	0.8972	0.8120
	ResNet-50	0.9118	0.9615	0.9036	0.8037
Normal and Pneumonia after data redistribution	VGG-16	0.9735	0.9834	0.9800	0.9480
	ResNet-50	0.9633	0.9834	0.9663	0.9113
	SqueezeNet	0.9667	0.9834	0.9708	0.9235
	Ensemble	0.9787	0.9941	0.9767	0.9388
Other virus/bacteria Pneumonia and COVID-19	VGG-16	0.9788	1	0.8966	0.9740
	SqueezeNet	0.9859	0.9615	0.9615	0.9913
	Ensemble	0.9859	1	0.9286	0.9837

Table 1: Overview of our result in Phase 1 and Phase 2

to constrain the model to make sure the translation from one domain to the other and back again to the original domain.

We implemented a 7×7 PatchGAN[21] discriminator that specially penalizes the scheme at the scale of patches. This discriminator can classify whether each patch of an input is real or fake. While the generator has an encoder-decoder model architecture. It will firstly down-sample the input image to a bottleneck layer, then interpret the encoding results with several ResNet layers to skip connections followed by a number of layers that decode the representation to the size of the target output images.

After training our CycleGAN model, we only make use of $G_{X \to Y}$ to observe the translation effect on our test scans.

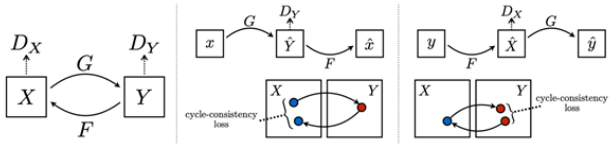


Figure 7: CycleGAN Architecture

5. Experiments

5.1. Experiment Workflow

In phase 1, we tried VGG-16 and ResNet-50 on the original dataset before the redistribution of data. For VGG-16, we set learning rate as 1e-4 and batch size of 20 for 20 epochs. In the case of ResNet-50, we set batch size as 32 for 40 epochs. After we applied adjustment to the dataset, we trained SqueezeNet mentioned in last section with the same batch size and epochs as ResNet. Finally, we applied different ensemble strategies within the 3 models and found the optimal one according to the model evaluation.

In phase 2, VGG-16 was trained respectively at temporal stage with dataset only from[9], adding bacteria pneumonia scans from[22] mixing with virus pneumonia as other pneumonia class other than COVID-19, adding heterogeneous COVID-19 scans from diverse source collected in[11]. And in the last stage, we trained SqueezeNet again on this new dataset as it performed better than ResNet-50 in phase 1 and implemented model ensemble to see whether there could be an improvement. All the hyperparameters were set to be the same as in phase 1 as the datasets share similar features.

In both of the phases, we chose Adam[23] as optimizer and adopted the method of early stop with validation accuracy as monitor and only saved the best model weights to perform prediction on test set.

In bone suppression task, we set batch size as 1 and stored $G_{X \to Y}$ and $G_{Y \to X}$ for every epoch as it took hours to train an epoch even on a relatively fast GPU. Moreover, we randomly sampled 5 of the original chest X-ray scans to observe the generated translation effect in the whole process. Finally, we ended the training after 13 epochs and obtained the final saved model to test on new inputs.

5.2. Model Evaluation

For phase 1 and phase 2, we will show our results based on test accuracy, sensitivity, precision and specificity. The overview of the performance of our model can be found in Table 1 and the confusion matrix of the ensemble models for both of the phases can be found in Figure 8.

It turns out that VGG-16 can achieve highest test accuracy after the data redistribution among the 3 models we chose in phase 1. And after model ensemble, the accuracy can be further improved from 97.35% to 97.87%. While in phase 2, SqueezeNet will outperform all the other models including the ensemble with highest accuracy 98.59% and comprehensively higher scores for other indicators. Though ResNet-50 gives us a relatively unsatisfactory result compared to other models in phase 1, we tried utilizing it in ensemble. It is shown under expectation that the ensemble

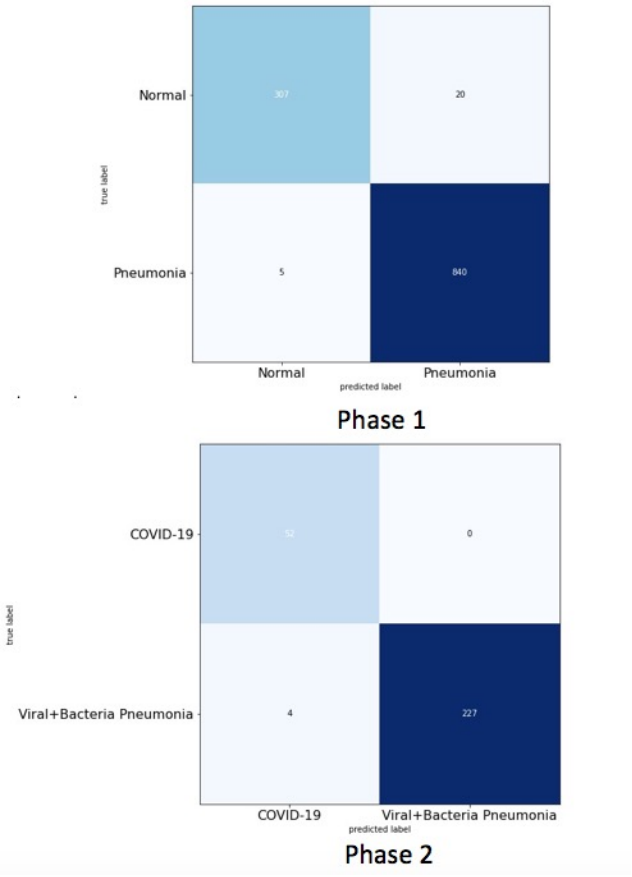


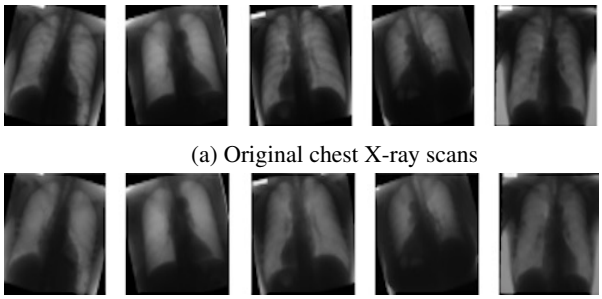
Figure 8: Confusion matrix of ensemble model

from all the models is not as good as the one solely using VGG-16 and SqueezeNet. Thus, we left out ResNet-50 in both of the phases. From Table 1, we can make a conclusion that our strategies of data redistribution and model ensemble do can boost the result. Assigning appropriately more data to validation can help the model monitor the loss and make a wise decision in a less random way, while ensemble can reduce the variance of a single model and make the model less sensitive to different initialization which also reduces the randomness[13].

For bone suppression task, we merely evaluated the results by its visual effect. And the bone-suppressed results generated from original chest X-ray scans can be seen in Figure 9.

6. Conclusion

In summary, we developed a useful ensembled CNN model for the classification task of pneumonia X-ray scans (Phase 1) and COVID-19 X-ray scans (Phase 2). Besides, we also proposed a possible solution for bone-suppression task: using CycleGAN to accomplish the goal of training a bone-suppress model, which has not been pointed out by other researchers before. Satisfactory results have been achieved in the tasks of classifying pneumonia scans from healthy scans and COVID-19 scans from pneumonia scans. The Accuracy, Sensitivity, Precision and Specificity are shown in Table 1, which has outperformed the state of the art classification models for COVID-19.



(a) Original chest X-ray scans

(b) Generated bone-suppressed results

Figure 9: Visual effect of bone-suppression from CycleGAN after 13 epochs

Transfer learning techniques and ensembled model techniques are the most important techniques we learned through the development of our ensembled models. Besides, we also learned about the application of CycleGAN in the medical field, which provides a novel idea of connecting GAN with the image processing techniques dealing with medical images.

Apart from the techniques we learnt about the development of models, we also gained knowledge about processing data and determining the classes. The performance of our model is quite limited in phase 1 originally, and it turns out to be that the original dataset does not have a reasonable partition of the training, validation and testing sets. The careful determination of classes is also an important act to take: Initially, we decided to classify the X-ray scans into COVID-19, healthy and non-COVID-19 classes, but the performance was poor comparing to the 'divide-and-conquer' two-phase model we used eventually. These remind us to pay attention to the data pre-processing as well as classification policy, which have been ignored by us before implementing this project.

As for the future extensions and novel applications, we are planning to extend our model to detect COVID-19 from both CT-scans and X-ray scans, instead of X-ray scans only. We believe that the decisions made from both CT-scans and Chest X-ray scans of the patients can be more accurate and reliable compared with our current model. Another possible extension is that we want to enhance our model in phase 2 to enable it to classify the actual types of non-COVID viral/bacterial scans. Currently, our model is only capable of diagnosing COVID-19 scans from viral/bacterial pneumo-

nia scans, which is not sufficient for the comprehensive applications in the hospitals. If our model in phase 2 can determine other common repository diseases such as: MERS[2], SARS[3] and ARDS[1], then it would have higher value and more possibilities to be applied in the real diagnosis process, which will make great contributions to the human society.

For the bone-suppress model we implemented using CycleGAN, real life user experiments and evaluations are needed to see if it can improve the performance of doctors to diagnose the repository diseases. We believe that it will be of help to the diagnosis process if properly utilized. Besides, we think that another possible use of this bone-suppress model is serving as a pre-processing stage for other classification tasks of specimen. Experiments and evaluations are needed for these extensions and applications.

References

[1] Acute respiratory distress syndrome, May 2020.

[2] Middle east respiratory syndrome-related coronavirus, May 2020.

[3] Severe acute respiratory syndrome, May 2020.

[4] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia. Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in china: a report of 1014 cases. *Radiology*, page 200642, 2020.

[5] A. Almahairi, S. Rajeswar, A. Sordoni, P. Bachman, and A. Courville. Augmented cyclegan: Learning many-to-many mappings from unpaired data. *arXiv preprint arXiv:1802.10151*, 2018.

[6] H. X. Bai, B. Hsieh, Z. Xiong, K. Halsey, J. W. Choi, T. M. L. Tran, I. Pan, L.-B. Shi, D.-C. Wang, J. Mei, et al. Performance of radiologists in differentiating covid-19 from viral pneumonia on chest ct. *Radiology*, page 200823, 2020.

[7] C. Butt, J. Gill, D. Chun, and B. A. Babu. Deep learning system to screen coronavirus disease 2019 pneumonia. *Applied Intelligence*, page 1, 2020.

[8] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in wuhan, china: a descriptive study. *The Lancet*, 395(10223):507–513, 2020.

[9] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al-Emadi, et al. Can ai help in screening viral and covid-19 pneumonia? *arXiv preprint arXiv:2003.13145*, 2020.

[10] C. Chu, A. Zhmoginov, and M. Sandler. Cyclegan, a master of steganography. *arXiv preprint arXiv:1712.02950*, 2017.

[11] J. P. Cohen, P. Morrison, and L. Dao. Covid-19 image data collection. *arXiv 2003.11597*, 2020.

[12] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[13] T. Dietterich. Ensemble methods in machine learning. pages 1–15, 01 2000.

[14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 2672–2680. Curran Associates, Inc., 2014.

[15] O. Gozes, M. Frid-Adar, H. Greenspan, P. D. Browning, H. Zhang, W. Ji, A. Bernheim, and E. Siegel. Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis. *arXiv preprint arXiv:2003.05037*, 2020.

[16] M. Gusarev, R. Kuleev, A. Khan, A. R. Rivera, and A. M. Khattak. Deep learning models for bone suppression in chest radiographs. In *2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–7. IEEE, 2017.

[17] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[18] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[19] M.-C. Huynh. X-ray bone shadow supression, Oct 2018.

[20] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size. *arXiv preprint arXiv:1602.07360*, 2016.

[21] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks, 2016.

[22] D. Kermany, K. Zhang, and M. Goldbaum. Labeled optical coherence tomography (oct) and chest x-ray images for classification. *Mendeley data*, 2, 2018.

[23] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization, 2014.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[25] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, et al. Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest ct. *Radiology*, page 200905, 2020.

[26] M.-Y. Ng, E. Y. Lee, J. Yang, F. Yang, X. Li, H. Wang, M. M.-s. Lui, C. S.-Y. Lo, B. Leung, P.-L. Khong, et al. Imaging profile of the covid-19 infection: radiologic findings and literature review. *Radiology: Cardiothoracic Imaging*, 2(1):e200034, 2020.

[27] W. H. Organization et al. Coronavirus disease 2019 (covid-19): situation report, 116. 2020.

[28] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition, 2014.

[29] L. Wang and A. Wong. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images. *arXiv preprint arXiv:2003.09871*, 2020.

[30] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, and W. Tan. Detection of sars-cov-2 in different types of clinical specimens. *Jama*, 2020.

[31] Y. Wang, H. Kang, X. Liu, and Z. Tong. Combination of rt-qpcr testing and clinical features for diagnosis of covid-19 facilitates management of sars-cov-2 outbreak. *Journal of medical virology*, 2020.

[32] T. Zhou, P. Krähenbühl, M. Aubry, Q. Huang, and A. A. Efros. Learning dense correspondence via 3d-guided cycle consistency, 2016.