Deep learning to distinguish COVID-19 from other lung infections, pleural diseases, and lung tumors

Ali Serener
Dept. of Electrical and Electronic Engr.
Near East University
Nicosia, Cyprus
ali.serener@neu.edu.tr

Sertan Serte
Dept. of Electrical and Electronic Engr.
Near East University
Nicosia, Cyprus
sertan.serte@neu.edu.tr

***Abstract*—COVID-19 is a highly infectious respiratory disease caused by severe acute respiratory syndrome coronavirus 2. It can lead to cough and fever and in some cases severe pneumonia. It is generally detected by reverse-transcription polymerase chain reaction and computed tomography scans. However, as it is a lung disease, it has common symptoms with other respiratory diseases. This necessitates us to carefully differentiate COVID-19 from such diseases during the diagnosis. This work aims to do that with the help of several deep learning architectures and chest radiographs. It specifically focuses on differentiating COVID-19 from pneumonia, pleural effusion and lung mass. During this analysis, it is shown that we can differentiate COVID-19 from other respiratory diseases using various deep learning architectures. It is further shown that ResNet-18 architecture produces the best overall performance in three scenarios of experiments.**

***Index Terms*—Chest radiographs, COVID-19, deep learning, lung mass, pleural effusion, pneumonia**

I. Introduction

Coronoviruses, named after their crown-shaped tips, are viruses that are known to cause respiratory diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). A never before seen coronovirus disease, named coronovirus disease 2019 (COVID-19), was reported for the first time in Wuhan, China, in December 2019. Being highly contagious, the disease has rapidly widespread worldwide, resulting in an epidemic.

With symptoms like fever, cough, myalgia, and dyspnea, COVID-19 has been mainly diagnosed using reverse-transcription polymerase chain reaction (RT-PCR). Besides RT-PCR, which is known to take a minimum of a day or two to produce results, doctors have used chest radiographs (CXR) and computed tomography (CT) to diagnose patients with COVID-19. However, medical imaging may lead to false-positives due to common characteristics between COVID-19 and other respiratory diseases such as pneumonia. Microbiologists and radiologists also have enormous pressure on them as they are working around the clock to produce the diagnosis results for currently millions of cases of COVID-19. This, in turn, makes hiring microbiologists and radiologists a high priority for the hospital administrators in hardest hit cities around the world.

Researchers have, therefore, used deep learning methods and medical images for the detection of various diseases around the body, such as those over the eye [1], [2] and the skin [3], [4]. Therefore, it is no surprising that with the recent surge of coronavirus cases, the researchers have focused on detecting COVID-19 also using such methods (see Section II)

This paper aims to use several deep learning architectures and medical images to clearly distinguish COVID-19 from other respiratory diseases that do have similar symptoms. Specifically, it is the goal of this paper to differentiate COVID-19 from pneumonia, pleural effusion, and lung mass (see Figure 1 for sample radiographs of these diseases). To accomplish this goal, we used six different deep learning architectures, ran COVID-19 detection experiments and analyzed the results. We believe, this may help eliminate the false-positives that may occur during COVID-19 diagnosis.

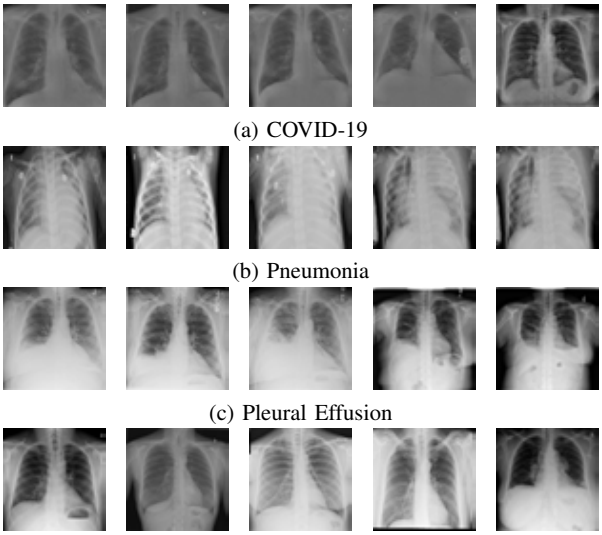
The organization of this paper is section II reviews the recent advances in deep learning detection of COVID-19 on medical images. Preprocessing and deep learning architecture details are briefly given in Section III. Section IV gives the hardware details. Section V summarizes the details of the four datasets used during the experiments. Section VI lists the performance results. These results are analyzed in Section VII. Section VIII overviews the conclusions drawn from the experiments.

II. Related Work

There have been several recent studies on the detection of COVID-19 using deep learning. In this section, we will summarize some of them.

Butt et al. [5] used deep learning models to classify CT chest scans for COVID-19, viral pneumonia or no infection and compared their detection rates to RT-PCR. In another work, Brunese et al. [6] proposed the use deep learning to detect COVID-19 from X-rays in three steps. In [7], Ardakani et al. used ten well-known convolutional neural networks to distinguish COVID-19 disease infection from non-COVID-19 diseases. Hu et al. [8] proposed a weakly supervised deep learning method to detect COVID-19 and distinguish it from non-COVID-19 diseases. In [9], Mei et al. included clinical symptoms, exposure history, and laboratory testing in the deep learning detection of COVID-19. Apostolopoulos [10] detected coronavirus disease, bacterial pneumonia and healthy cases from X-ray images using convolutional neural

978-1-7281-8073-1/20/$31.00 ©2020 IEEE



(a) COVID-19

(b) Pneumonia

(c) Pleural Effusion

(d) Lung Mass

Fig. 1. Chest radiographs showing four respiratory diseases.

networks. In [11], Pereira et al. detected COVID-19 resulted pneumonia from chest X-ray images using flat and hierarchical classification methods. Wang et al. [12] used a deep learning system and lung CT images for diagnosis and prognosis of COVID-19. Li et al. [13] developed a deep learning model to detect COVID-19, community acquired pneumonia, and non-pneumonia diseases from chest CT scans.

III. Method

A. Preprocessing

The images included in the datasets have been resized to 224×224×3.

B. Classification

To distinguish COVID-19 from pleural diseases, lung infections and lung tumors, we have used six different deep learning architectures, namely AlexNet, ResNet-18, ResNet-50, VGG, MobileNet-v2, and DenseNet-121. These architectures have eight, eighteen, fifty, sixteen, eight, and hundred and twenty one layers, respectively. We have pre-trained them on ImageNet dataset which contains approximately one million images.

IV. Implementation

We trained the deep learning architectures by a single NVIDIA GeForce GTX 1080Ti GPU with Caffe deep learning framework.

V. Datasets

To distinguish COVID-19 from other respiratory diseases, we have used the public datasets of Kermany et al. [14], Wang et al. [15], and Cohen [16].

The dataset of Kermany et al. [14] has 3883 images of chest X-rays containing pneumonia. It also contains 1349 chest X-ray images that do not contain any pneumonia. The dataset of [15] contains 14 classes of thoracic disease images. These categories include lung mass and pleural effusion. The dataset of Cohen [16] has chest X-ray images containing COVID-19.

The number of images used from each dataset is listed in Table I.

TABLE I
Number of images used.

Disease	Dataset	Train	Test
Pneumonia	Kermany et al. [14]	168	20
Pleural Effusion	Wang et al. [15]	168	20
Lung Mass	Wang et al. [15]	168	20
COVID-19	Cohen [16]	168	20
Total		672	80

We have also rotated the images we have used from each dataset to get the augmented images. The number of augmented images are listed in Table II.

TABLE II
Number of augmented images used.

Disease	Dataset	Train	Test
Pneumonia	Kermany et al. [14]	501	20
Pleural Effusion	Wang et al. [15]	501	20
Lung Mass	Wang et al. [15]	501	20
COVID-19	Cohen [16]	501	20
Total		2004	80

VI. Performance Evaluation

A. Distinguishing COVID-19 from pleural effusion

In this section, we will report on the classification performance of five deep learning methods to distinguish between COVID-19 and pleural effusion. To do that, we used 168 chest X-ray images of patients containing COVID-19 and pleural effusion for training and 20 chest X-ray images of patients containing COVID-19 and pleural effusion for testing. We obtained these images from the datasets of [15] and [16].

The performance results are as follows: The accuracy of MobileNet-v2 model is 0.75, its sensitivity is 1.00, and its specificity is 0.67. The area under the region of convergence (ROC) curve (AUC) is 0.78. For AlexNet model, the AUC is 0.74, the accuracy is 0.93, the sensitivity is 1.00 and the specificity is 0.87. For ResNet-18 model, the AUC is 0.67, the accuracy is 0.78, the sensitivity is 0.97 and the specificity is 0.69. The accuracy of ResNet-50 model is 0.79, its sensitivity is 0.96, and its specificity is 0.71. The AUC is 0.60. For VGG model, the AUC is 0.55, the accuracy is 0.79, the sensitivity is 0.95 and the specificity is 0.71 (see Table III).

Figure 2 is a performance comparison of the five architectures used to differentiate COVID-19 from pleural effusion. We observe from this figure that out of five models, MobileNet-v2 and AlexNet have performed the best.

TABLE III
DISTINGUISHING COVID-19 FROM PLEURAL EFFUSION.

Network	AUC	Accuracy	Sensitivity	Specificity
MobileNet-v2	0.78	0.75	1.00	0.67
AlexNet	0.74	0.93	1.00	0.87
ResNet-18	0.67	0.78	0.97	0.69
ResNet-50	0.60	0.79	0.96	0.71
VGG	0.55	0.79	0.95	0.71

TABLE IV
DISTINGUISHING COVID-19 FROM PNEUMONIA.

Network	AUC	Accuracy	Sensitivity	Specificity
ResNet-18	0.94	0.97	1.00	0.84
AlexNet	0.92	0.93	1.00	0.87
VGG	0.88	0.95	1.00	0.92
MobileNet-v2	0.84	0.96	1.00	0.93
ResNet-50	0.82	0.97	1.00	0.95

TABLE V
DISTINGUISHING COVID-19 FROM LUNG MASS.

Network	AUC	Accuracy	Sensitivity	Specificity
DenseNet-121	0.77	0.75	0.61	0.78
ResNet-18	0.65	0.76	0.65	0.79
VGG	0.64	0.75	0.62	0.79
MobileNet-v2	0.61	0.74	0.58	0.78
AlexNet	0.52	0.66	0.41	0.74
ResNet-50	0.30	0.75	0.59	0.80

B. Distinguishing COVID-19 from pneumonia

We have also evaluated distinguishing COVID-19 from pneumonia using five of the aforementioned architectures. For this performance analysis, we used from the dataset of [14] and [16] 168 chest X-ray images containing COVID-19 and pneumonia for training and 20 chest X-ray images containing COVID-19 and pneumonia for testing.

As listed in Table IV, the performance results are as follows: The accuracy of ResNet-18 model here is 0.97, its sensitivity is 1.00 and its specificity is 0.84. The AUC is 0.94. The accuracy of AlexNet model is 0.93, its sensitivity is 1.00, and its specificity is 0.87. The AUC is 0.92. For VGG model, the AUC is 0.88, the accuracy is 0.95, the sensitivity is 1.00 and the specificity is 0.92. For MobileNet-v2 model, the AUC is 0.84, the accuracy is 0.96, the sensitivity is 1.00 and the specificity is 0.93. The accuracy of ResNet-50 model is 0.97, its sensitivity is 1.00, and its specificity is 0.95. The AUC is 0.82.

Figure 3 shows the ROC curve comparison of the five architectures for differentiating COVID-19 from pneumonia. The analysis of this figure shows AlexNet and ResNet-18 models have performed the best.

C. Distinguishing COVID-19 from lung mass

Next, we would like to report the performance results observed in differentiating COVID-19 from lung mass. From

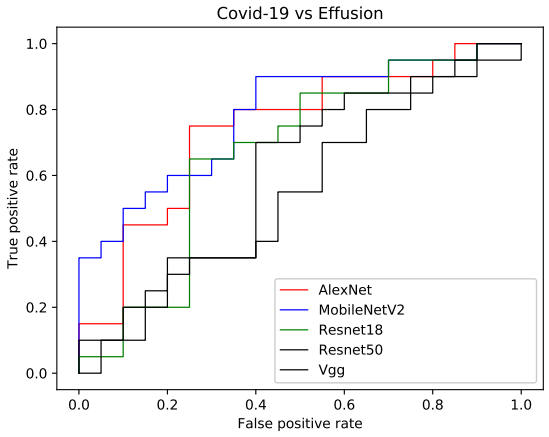


Fig. 2. Comparison of the ROC curves for distinguishing COVID-19 from pleural effusion.

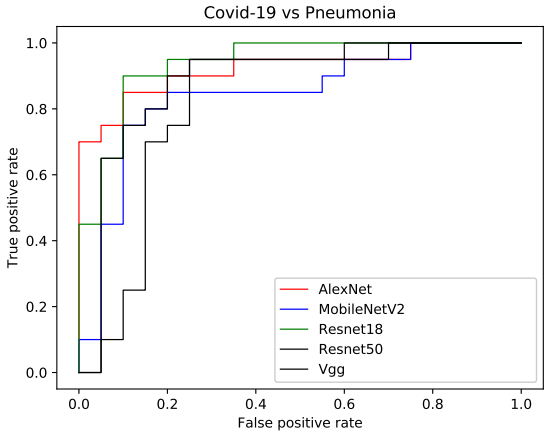


Fig. 3. Comparison of the ROC curves for distinguishing COVID-19 from pneumonia.

the datasets of [15] and [16], we used 168 chest X-ray images of patients having COVID-19 and lung mass for training and 20 images having COVID-19 and lung mass for testing.

The performance results are listed in Table V. These results are as follows: the accuracy of DenseNet-121 classification is 0.75, its sensitivity is 0.61 and its specificity is 0.78. The AUC is 0.77. For ResNet-18 model, the AUC is 0.65, the accuracy is 0.76, the sensitivity is 0.65 and the specificity is 0.79. The AUC of VGG model is 0.64, the accuracy is 0.75, the sensitivity is 0.62 and the specificity is 0.79. For MobileNet-v2 model, the AUC is 0.61, the accuracy is 0.74, the sensitivity is 0.58 and the specificity is 0.78. The accuracy of AlexNet classification is 0.66, its sensitivity is 0.41 and its specificity is 0.74. The AUC is 0.52. Finally, the AUC of ResNet-50 model is 0.30, the accuracy is 0.75, the sensitivity is 0.59 and the specificity is 0.80.

Figure 4 displays a performance comparison of the six

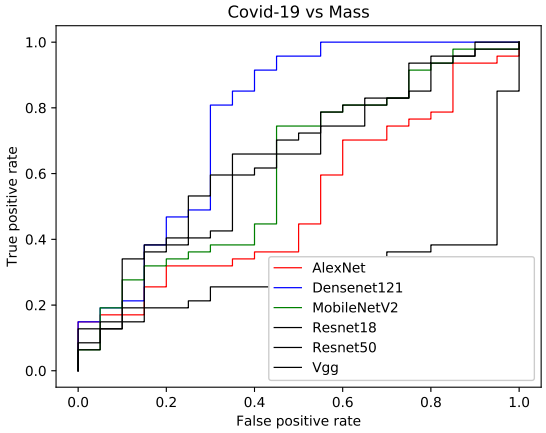


Fig. 4. Comparison of the ROC curves for distinguishing COVID-19 from lung mass.

architectures used to differentiate COVID-19 from lung mass. The analysis of this figure provides that DenseNet-121 model is the best in distinguishing COVID-19 from lung mass.

VII. Discussions

To separate COVID-19 from other lung infections, pleural diseases and lung tumors, we used deep learning methods on chest radiographs. Specifically, we used six methods to distinguish COVID-19 from pneumonia, pleural effusion and lung mass. Using the datasets of [14], [15], and [16], we trained 169 chest radiograph images having the four types of respiratory diseases. We then used MobileNet-v2, AlexNet, ResNet-18, ResNet-50, VGG, and DenseNet-121 architectures to test each of the 20 chest radiograph images for distinguishing COVID-19 from each of the three diseases. We have listed the performance results of each deep learning architecture on the three cases of experiments. These results show that in distinguishing COVID-19 from pleural effusion, MobileNet-v2 and AlexNet architectures have performed the best. They also show that we can distinguish COVID-19 from pneumonia best using ResNet-18 and AlexNet architectures. They further show that we can separate COVID-19 from lung mass the best using DenseNet-121 and ResNet-18 architectures. Out of the six deep learning architectures, it is the ResNet-18 architecture that is shown to perform the best in separating COVID-19 from pleural effusion, pneumonia, as well as lung mass.

VIII. Conclusions

In this paper, we performed three cases of experiments to analyze how we can best distinguish COVID-19 from other lung infections, pleural diseases, and lung tumors using several deep learning methods. What we specifically wanted to focus on was how COVID-19 can be best distinguished from pneumonia, pleural effusion and lung mass. We tabulated the performance results in terms of AUC, accuracy, sensitivity, and specificity metrics. The analysis of these results indicate that we can distinguish COVID-19 from each of the three diseases using different deep learning architectures. However, ResNet-18 architecture has given the best overall performance results if we consider all three cases of experiments.

References

[1] S. Kaymak and A. Serener, "Automated age-related macular degeneration and diabetic macular edema detection on oct images using deep learning," in *IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP)*. IEEE, 2018, pp. 265–269.

[2] A. Serener and S. Serte, "Transfer learning for early and advanced glaucoma detection with convolutional neural networks," in *2019 Medical Technologies Congress (TIPTEKNO)*. IEEE, 2019, pp. 1–4.

[3] S. Kaymak, P. Esmaili, and A. Serener, "Deep learning for two-step classification of malignant pigmented skin lesions," in *14th Symposium on Neural Networks and Applications (NEUREL)*. IEEE, 2018, pp. 1–6.

[4] A. Serener and S. Serte, "Keratinocyte carcinoma detection via convolutional neural networks," in *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 2019, pp. 1–5.

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

[6] L. Brunese, F. Mercaldo, A. Reginelli, and A. Santone, "Explainable deep learning for pulmonary disease and coronavirus covid-19 detection from x-rays," *Computer Methods and Programs in Biomedicine*, p. 105608, 2020.

[7] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "Application of deep learning technique to manage covid-19 in routine clinical practice using ct images: Results of 10 convolutional neural networks," *Computers in Biology and Medicine*, p. 103795, 2020.

[8] S. Hu, Y. Gao, Z. Niu, Y. Jiang, L. Li, X. Xiao, M. Wang, E. F. Fang, W. Menpes-Smith, J. Xia *et al.*, "Weakly supervised deep learning for covid-19 infection detection and classification from ct images," *IEEE Access*, 2020.

[9] X. Mei, H.-C. Lee, K.-y. Diao, M. Huang, B. Lin, C. Liu, Z. Xie, Y. Ma, P. M. Robson, M. Chung *et al.*, "Artificial intelligence–enabled rapid diagnosis of patients with covid-19," *Nature Medicine*, pp. 1–5, 2020.

[10] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, p. 1, 2020.

[11] R. M. Pereira, D. Bertolini, L. O. Teixeira, C. N. Silla Jr, and Y. M. Costa, "Covid-19 identification in chest x-ray images on flat and hierarchical classification scenarios," *Computer Methods and Programs in Biomedicine*, p. 105532, 2020.

[12] S. Wang, Y. Zha, W. Li, Q. Wu, X. Li, M. Niu, M. Wang, X. Qiu, H. Li, H. Yu *et al.*, "A fully automatic deep learning system for covid-19 diagnostic and prognostic analysis," *European Respiratory Journal*, 2020.

[13] 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*, p. 200905, 2020.

[14] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[15] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2097–2106.

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