A Novel Approach to Detect and Monitor COVID-19 Infection Using Transfer Learning Concept in AI

Vinh Pham[1], Ha Min Son[1], Thuy Huynh[2], and Tai-Myoung Chung[1(✉)]

[1] Department of Computer Science and Engineering, Sungkyunkwan University,
Suwon, Republic of Korea
{vinhpham,sonhamin}@g.skku.edu, tmchung@skku.edu
[2] Institute for Quantum Biophysics (IQB) and Department of Biophysics,
Sungkyunkwan University, Suwon, Republic of Korea

Abstract. COVID-19 is an ongoing pandemic that is continuing to spread after recording one hundred million cases, causing millions of casualties, overwhelming health care systems of many countries, and threatening the whole world. Monitoring and assessing the severity of COVID-19 through artificial intelligence would be a practical support for medical practitioners reviving patients and offloading the burden from medical system. Previous works exploited deep learning, for this purpose, which produces inexplainable diagnosis results and lacks medical evidence. Integrating clinical symptom into diagnosis with deep learning will support generating results more compelled and validated. In this study, we focus on verifying the effectiveness of applying the human lung lesion, specifically Ground Glass Opacity and Consolidation, caused by typical pneumonia for COVID-19 detection or severity assessment on chest X-ray image with deep learning technology. We have conducted multiple experiments with state-of-art machine learning architectures (MobileNetV2, ResNet, Faster R-CNN) on many datasets to establish the conclusion. The experiment result demonstrates that lung lesion is useful when incorporating with deep learning solutions for monitoring COVID-19 progression and will provide solid pathway to develop an improved model and support better research in the future.

Keywords: COVID-19 · Deep learning · Severity assessment

1 Introduction

The COVID-19 pandemic is an ongoing pandemic of coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARSCoV-2). Recent studies reported that the pathologic mechanism of COVID-19 is similar to other severe acute respiratory coronavirus pneumonia. In this research, we utilize features that are typically associated with pneumonia: Ground Glass Opacity (GGO) and Consolidation, to assess the severity of the COVID-19 case of concern rather than detection of the case itself. Due to the lack of credibility of Machine-Learning (ML) based solutions,

this research focuses on the detection of GGO and Consolidation from the X-ray of COVID-19 patients and measures the size and position this to evaluate the severity of the given case of COVID-19. Convolutional neural network (CNN) has demonstrated its effectiveness in diagnosing medical radiography (e.g., chest Xray, CXR and computed tomography, CT) and perfectly matches this work's purpose.

In this study, first, we demonstrate the usefulness of lung lesions (GGO and Consolidation) as a bio-marker in developing an AI-powered solution for COVID-19 diagnosis and severity assessment. Next, we detect the lung lesion on CXR by the most current state-of-the-art deep learning architecture model (e.g., Faster R-CNN) for object detection. Moreover, the standalone employment of the feature extractor of Faster R-CNN as an image classifier for conducting experiments, is our premise for applying transfer learning in proposals to come.

Despite the limitations of this study, which are related to the capability of annotating sufficient amounts of training data, results concerning the exploitation of pneumonia features are encouraging. Especially, GGO and Consolidation incorporated with deep learning facilitates more specialized research into COVID-19 diagnosis and severity assessment on the development of medical radiography solutions, which will contribute to the fighting against the pandemic.

2   Related Work

There was a markedly limited number of existing works on the quantitative monitoring of COVID-19 progression or severity assessment [1]. The majority of these existing studies focus only on computed tomography (CT) modality. However, CT image data acquisition requires complex patient preparation, extensive expertise, time consuming procedures, and the use of highly sophisticated devices (e.g., Picture archiving and communication systems, PACS). On the contrary, Chest X-ray (CXR) is the first-line image modality recommended by medical practitioners for lung infection detection. CXR may have the potential to play a critical role in COVID-19 related problems.

Although there was some notable research on automated (thanks to deep learning technology) severity assessment for typical pneumonia or especially COVID-19 [2] using CXR, these studies fail to explain the decisions made by their proposed models in terms of medical evidence. For instance, when assessing the level of severity, the presence of lung lesions in both the left and right lung (bilateral distribution) should be considered more severe than cases with lung lesion presence in just one lung. Medical evidence, as the basis of accurate diagnosis and decision making, is a crucial requirement in healthcare applications, especially COVID-19 severity assessment. However, it is difficult to explain results produced by AI-powered solutions in which convincing medical evidence is absent.

These unsatisfactory results from previous studies motivated us to explore the usefulness of lung lesions in developing an AI based solution for COVID-19 as well as poses a non-trivial task for this research that is effectively detecting lung lesion on CXR with their properties (i.e. size, position, confidence, etc.).

3   Methodology and Material

We have conducted multiple experiments with numerous machine learning architectures on multiple datasets in this research. In this section, we will first discuss the connection between the COVID-19 and pneumonia lung lesions. Next, we will introduce the employed public datasets and manipulate them to become input data for the experiments. Finally, we will explain the employed machine learning architectures that were adjusted to the demands of different experiments.

3.1   COVID-19 and Lung Lesion

According to medical reports [3], COVID-19 is categorized as atypical pneumonia (i.e., severe acute respiratory system SARS or Middle East respiratory syndrome MERS). Apparently, lung injuries caused by COVID-19 infections are similar to those caused by another coronavirus. Chest image studies [4] show that the two types of lung lesions that are most frequently observed in COVID-19 patients are GGO and Consolidation. The level of whiteness in the lungs, on radiography, depends on the severity of the pneumonia case. The covering of increased whiteness partially obscuring lung markings on radiography signifies the existence of GGO. When lung markings are completely lost in whiteness, it means that GGO has developed into Consolidation [3]. Figure 1 shows these two lesions on CXR.
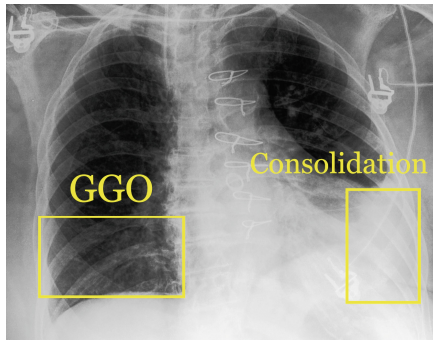

Fig. 1. GGO and consolidation.

3.2   Dataset

The CheXpert dataset [5] consists of 224,316 chest radiographs collected from Stanford Hospital belonging to 65,240 patients. Board-certified radiologists have annotated 14 observations (e.g., GGO, Edema, Consolidation, etc.) on these radiographs and supplied them as the metadata (CSV format). With this metadata, we can filter all images that contain GGO and Consolidation. These filtered images will be used as training data for training one model as a binary classifier to detect between GGO versus Consolidation.

However, an issue occurred when we intended to use this data to train a machine learning model for detecting lung lesions. Because the lesion was annotated for the entire image, it was not possible to locate the lesion's properties within the CXR. Thanks to the collaboration of a medical doctor and a radiologist, we selected 31 CXRs that contain the most common patterns of both GGO and Consolidation, annotated them (with a bounding box) and stored them in the COCO object annotation standard format for the training process. Table 1 displays the statistics of the lung lesion data we prepared. Because one CXR may contain multiple lesions, the Annotation(s) column displays the total annotated instance for each lesion type.

Table 1. Lung lesion data statistics.

Task	Lesion	Image(s)	Annotation(s)
Classification	GGO	3345	
	Consolidation	791	
Detection	GGO	31	56
	Consolidation	31	37

We employed the COVID-Xray-5k dataset from [6] for training another model as a binary classifier to classify COVID-19 and non-COVID-19 CXRs. The test data of COVID-Xray-5k was seriously unbalanced, which could possibly lead to inconsistent results if not rectified. Therefore, we had to re-balance the test data of COVID-Xray-5k by decreasing the number of non-COVID-19 CXR and then updating the COVID-19 CXRs using images from a public COVID-19 Image Data Collection dataset [7]. Table 2 displays the number of CXRs in the training dataset as well as the testing dataset before and after balancing the COVID-Xray-5k dataset.

In order to verify whether the above-mentioned types of lung lesions have a correlation to COVID-19, the lung lesion classifier will be evaluated using the COVID-19 testing dataset and compare with the COVID-19 classifier's performance (i.e., cross checking).

Table 2. COVID data statistics.

Dataset		Train	Test
COVID-Xray-5k	COVID	84	100
	Non-COVID	2000	3000
COVID-Xray-5k balanced	COVID	84	**244**
	Non-COVID	2000	**300**

3.3   Machine Learning Architecture

Faster R-CNN: is currently the state-of-the-art model for object detection or segmentation [8], Fig. 2 shows the overview architecture of Faster R-CNN, it consists of 3 main components, each performing a unique function:

– The *Backbone Network* extracts multiple-scale feature maps from the input images and facilitates object detection at different scales.
– The *Region Proposal Network* performs its algorithm on extracted feature maps and generate locations of possible objects in the input image.
– The *ROI (region of interest) Head* will classify the object and predict its location based on the feature maps and proposed locations.
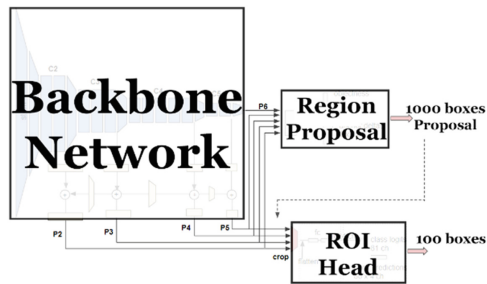

Fig. 2. Faster R-CNN architecture.

A unique feature of Faster R-CNN is that the Backbone Network can be constructed along with popular and preferred CNN architectures (e.g., ResNet and MobileNet). In consideration of the time-consuming and expert involvement required in the lung lesion annotating process, we hypothesize that we should apply the transfer learning technique and fine-tuning strategy for training the model in order to be able to deploy the proposed solution in the multi-development phase. The idea was to first train only the *Backbone Network* of the Faster R-CNN as an image classifier for classifying lung lesions using roughly labeled data, i.e., lung lesions annotated for the entire image. Due to the abundant amount of lung lesion labeled (on entire image) data available, training the *Backbone Network* first as an image classifier will help it sufficiently capture the lesion's characteristics. In urgent situations induced by the pandemic, this classifier may be deployed as an AI-powered lung lesion detection utility that provides overburdened medical professionals a second pair of eyes that support pneumonia COVID-19 diagnosis.

At the same time, we will annotate lung lesion properties. Because this process requires much time and medical expertise, we may only be able to produce a small volume of data compared to the lung lesion images without property labels. After acquiring a sufficient amount of annotated data, we will fine-tune the trained *Backbone Network* with the two remaining components of Faster R-CNN (i.e. *Region Proposal Network* and *ROI*). This process will help the trained *Backbone Network* evolve into a lung

lesion property detector. The knowledge gained from previous training for lung lesion classification tasks will be transferred and fine-tuned to the lesion properties detection task by later annotated limited size data.

In summary, we train two classifiers for COVID-19 and lung lesions using Google's MobileNetV2 [9]. which introduced new architectures of the Convolutional Neural Network (CNN) to improve its performance in mobile and resource-constrained environments. This allows us to deploy our model in embedded design applications. Also, we use Faster R-CNN with the ResNet-50 as the Backbone Network to train the lung lesion properties detector.

4 Experiment and Result

4.1 Training Procedure

In order to get the model (i.e., neural networks) better generalization to unseen data (i.e., evaluating with test data or inferencing realistic data), we preferred the method of training the model in batches of data. The iteration term refers to the total number of times a batch went through the model.

We divided the dataset into multiple 20-CXR batches in order to train two classifier models and fed them to the model for 200,000 iterations. In the case of the lung lesion property detector, due to the small size of the dataset used, we partitioned them into 2-CXR batches (ten times smaller) and fed them into the model for 10,000 iterations.

We also monitored the loss value of these models over the training iterations. Figure 3 displays the monitored loss value over the training iteration of three models. In both cases, the loss value nearly approached zero. This means that the model converged to its optimal point. Proving the success of the training process.
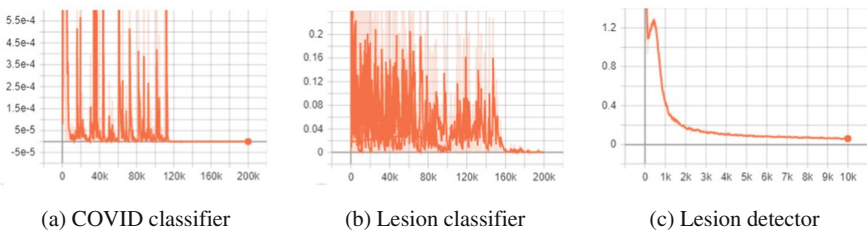

(a) COVID classifier   (b) Lesion classifier   (c) Lesion detector

Fig. 3. Monitored loss value over training iteration.

4.2 Result

After completing the training, we evaluate two classifiers with their own testing dataset. The COVID-19 classifier was able to achieve a Top-1 accuracy of up to 96.77%. Compared to the claimed results of [6] which are usually over 90% using the same COVID-19 testing dataset, we can confirm the correctness of our experiment implementation. The lung lesion classifier only achieves 62.83% Top-1 accuracy. Although this result does

not seem impressive, it still achieves better diagnosis compared to a random guess (i.e., accuracy better 50%), which means that the model have learned distinct features of GGO and Consolidation lesion. It appears that an increased amount of optimizing and training data is required to improve the lung lesion classifier's performance.

When we evaluate the COVID-19 classifier using the updated and balanced testing dataset, its performance accuracy dropped to 55.14%. This behavior shows us that it seems there is no specific radiography feature that can be exploited for detecting COVID-19.

Next, we performed the cross-check by evaluating the lung lesion classifier, which was not trained with COVID-19 CXR data at any point in time, with the updated and balanced COVID-19 test data. The lung lesion classifier achieved a higher accuracy (61.39%) in the classification of COVID-19 versus non-COVID-19 images in comparison to the above COVID-19 classifier's performance (55.14%). This result helps us verify the correlation between lung lesions (i.e., GGO and Consolidation) and lung injury caused by COVID-19 infection. Table. 4 displays the summarized results of both classifiers on their testing dataset as well as updated and balanced COVID-19 test data.

Table 4. COVID data statistic.

Test data	Model	Top-1 accuracy (%)
COVID-Xray-5k	COVID classifier	96.77
CheXpert	Lung lesion classifier	62.83
COVID-Xray-5k balanced	COVID classifier	55.14
	Lung lesion classifier	61.39

As is apparently demonstrated by the above experiments, lung lesion GGO and Consolidation have contributed to better detection of COVID-19. However, the result is not qualified in terms of sensitivity and specificity and should not be established as an independent criterion by its own. Instead, incorporating GGO and Consolidation with AI technology for COVID-19 severity assessment may be the wiser and more practical approach. This exploration motivates us to construct the lung lesion (along with its properties) detector using CXR to support COVID-19 severity assessment in future work.

Due to the limitation of lung lesion property annotated data, we were not able to conduct the detector model's performance evaluation. However, we inferred a sequence of CXRs with an interval of 3 days of a COVID-19 patient belongs to COVID-19 Image Data Collection dataset [7]. Figure 4 displays the GGO detection result with a bounding box and confidence score using our lung lesion property detector. The detector model can detect the GGO development of COVID-19 progression. Intuitively, this kind of result will contribute to the monitoring or assessment of the severity of given COVID-19 cases.

5   Conclusion

In this study, we have verified the correlations between two typical lung lesions of pneumonia, which are GGO and Consolidation, and COVID-19. Although incorporating lung lesion did help improve the detection accuracy, detecting COVID-19 based on radiography might be a less practical direction to move forward to.

We also demonstrated the capability of detecting GGO and Consolidation with properties by a detector model based on a Faster R-CNN architecture. The results are promising and has potential use in application for severity assessment or monitoring of COVID-19 progression, which are important factors in decision-making (e.g. determining whether to apply a mechanical ventilator, tracheal intubation, etc.).
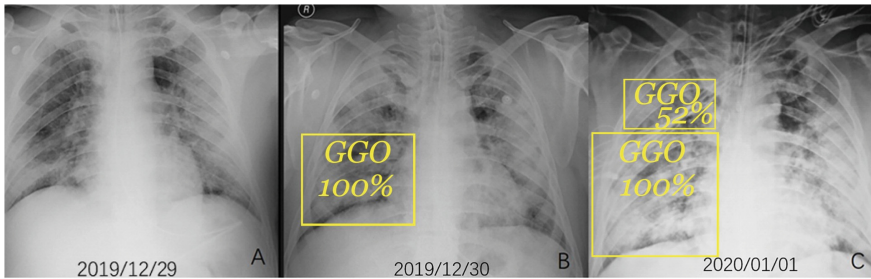

Fig. 4. Monitoring GGO size and position on COVID-19 CXR

References

1. Huang, L., et al.: Serial quantitative chest CT assessment of covid-19: a deep learning approach. Radiol.: Cardiothorac. Imaging **2**(2), e200075 (2020)
2. Signoroni, A., et al.: End-to-end learning for semiquantitative rating of covid-19 severity on chest x-rays. arXiv preprint arXiv:2006.04603 (2020)
3. Cleverley, J., Piper, J., Jones, M.M.: The role of chest radiography in confirming covid-19 pneumonia. bmj **370**, m2426 (2020)
4. Kanne, J.P.: Chest CT findings in 2019 novel coronavirus (2019-ncov) infections from Wuhan, China: key points for the radiologist (2020)
5. Irvin, J., et al.: Chexpert: a large chest radiograph dataset with uncertainty labels and expert comparison. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 590–597 (2019)
6. Minaee, S., Kafieh, R., Sonka, M., Yazdani, S., Soufi, G.J.: Deep-covid: predicting covid-19 from chest x-ray images using deep transfer learning. Med. Image Anal. **65**, 101794 (2020)

7. Cohen, J.P., Morrison, P., Dao, L., Roth, K., Duong, T.Q., Ghassemi, M.: Covid19 image data collection: prospective predictions are the future. arXiv preprint arXiv:2006.11988 (2020)
8. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. **39**(6), 1137–1149 (2016)
9. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: MobileNetV2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510–4520 (2018)