Detect COVID-19 with chest x-ray images using Deep learning

[1]Nguyen Hoang Minh Ngoc, [2]Nguyen Thi Tuyen Tran

[1]Student, [2]Medical Doctor
[1]Department of Mathematics,
[1]Le Quy Don High School for the Gifted, Khanh Hoa, VIETNAM

Abstract : COVID-19 was recognized as a pandemic by WHO on March 12, 2020, putting some countries in a situation of lacking resolve, capacity, and resources [27]. Inspired by these facts, deep learning algorithms relating to computer vision have been researched, and then applied to detect the COVID-19 using the chest X-ray by building a model called AiCOVID with an accuracy rate of 95.765%. Besides that, our model can detect whether COVID-19-inpatients have other lung diseases (viral pneumonia, lung cancer...). Notably, our model concerns fake positive cases of patients who had COVID-19 before.

Index Terms : COVID-19, computational biology, deep learning, computer vision, X-ray

I. INTRODUCTION

The earliest cases treated as unusual pneumonia were reported in Wuhan, the capital city of Hubei in China. The COVID-19 pandemic is now considered one of the deadliest epidemics, with the number of COVID-19 cases increasing. The COVID-19 epidemic has created havoc for the health and financial systems of the world [3, 12, 16, 18, 23, 25].

Despite governments and researchers' efforts, this virus continued to spread globally. However, the limitations of current diagnostic methods challenged countries to identify and manage the pandemic. Two commonly used viral tests are rapid diagnostic tests and reverse transcription polymerase chain reaction (RT-PCR), which is the gold-standard method [10, 24]. Some large-scale studies imply that a suspected case of COVID-19 requires multiple tests over the 14-day observation period to be considered as a true negative. This can make patients tense and can be a financial burden to hospitals [12, 13]. As a result, a method that can quickly and precisely identify patients is critical to reducing COVID-19 cases and providing patients with care to avoid complications.

Meanwhile, chest radiography is very useful in identifying patients in the intermediate and advanced stages of COVID-19 [2, 4, 6, 7, 11, 15, 17, 20, 22]. Some recent studies show that chest X-rays should be used to identify suspected COVID-19 patients thanks to its availability and low cross-infectivity risks when carried out [2, 19]. Besides that, artificial intelligence has been used to detect COVID-19 using chest x-rays of potential COVID-19 patients. CNNs, deep learning algorithms, are used to classify medical images as they have high accuracy rates [5, 8, 9].

II. METHODOLOGY

Study design

Our study began by collecting samples for four classes: COVID-19, normal, viral pneumonia, and lung cancer. After that, the samples are processed and divided into 5 sub-datasets on a random basis. A model was built on a Teachable Machine with CNN algorithms, using collected data sets.

Datasets collection

Samples were collected from prestigious and reliable resources [Appendix 2]. Especially in class Normal, there are samples from 19 cases which are recovered from COVID-19 patients and are followed up by doctors.

Then, the samples were converted into .png format and filtered to get the AP/PA images only. Finally, samples were named after the classes and numbered to make it easy for us to keep track of. Due to the limited storage space of our equipment, available samples were divided into 5 small data sets randomly. In each data set, 85% of the images are used for training purposes, and the rest are for evaluation and calculating other parameters.

Table 1. Data sets' components

Data set	COVID-19 class	Viral Pneumonia class	Lung cancer class	Normal class
1	301	300	247	300
2	264	300	247	300
3	308	308	247	300
4	308	297	247	300
5	319	308	247	341

Algorithm development

Our model's architecture is a convolutional neural network (CNN), consisting of three types of layers: convolutional, pooling, and fully-connected layers. While convolution and pooling layers perform feature extraction, fully connected layers map the extracted features into the final output.

There are some features that distinguish CNN from the other methods. First, CNN does not require human experts to extract features. Second, CNN architectures do not require hand-crafted segmentation of tumors or organs. [5]

Building the model

The model is built by Teachable Machine. After uploading the samples of each data set to 4 classes, the model is trained three times with different parameters, as follows:

Table 2. Training phases' properties

Phase	Epoch	Batch size	Learning rate
1st	100	16	0.001
2nd	50	16	0.001
3rd	50	16	0.0001

After done with a data set, all the samples will be replaced by the samples of the following data set until 5 data sets went through the model.

III. RESULTS

After being trained with 5 data sets, the model's accuracy is displayed in the following table:

Table 3. Model's accuracy

Class	Average accuracy	Lowest accuracy	Highest accuracy
Normal	0.9376	0.89	1
Lung cancer	0.998	0.93	1
COVID-19	0.945	0.9	0.98
Viral pneumonia	0.95	0.8	0.96

From the results, it can be inferred that the model's accuracy is both high and stable. Lung cancer has the highest average accuracy, followed by viral pneumonia and COVID-19 classes, and normal classes have the lowest accuracy. The average accuracy of four classes is 0.95765.

Our model is named AiCOVID. For further information about our model, please visit our GitHub site: https://github.com/angelinawong1210/AiCOVID

IV. DISCUSSIONS

Literature review

The fact that the RT-PCR tests are incapable of detecting COVID-19 at an early stage is a common problem encouraging the COVID-19 spread. An innovative diagnostic method is required to reduce the illness rate and allow crucial prevention steps. Thanks to the latest developments in convolutional neural networks, a potential approach using deep learning models is accessible. [7]

In a recent paper, Fang et al. conducted research about the sensitivity of lung CTs and RT-PCR tests in detecting COVID-19. The results showed that lung CTs are much more sensitive than RT-PCR tests on the first day (98% and 71%, respectively). [21]

Although CT images were initially used to identify suspected COVID-19 patients, some recent studies have shown that CXR images are preferred to CT images due to their popular availability and low cross-infectivity rate. The CXR scanners outweigh CT scanners as the machines' surfaces are easy to clean and there is no need to bring the patients into the radiography rooms. [19]

Comparisons with other models

Deep learning models detecting COVID-19 with chest X-ray have been built by laboratories and doctors around the world. Although there is no conflict between these, there were some differences in accuracy between models. [14]

Table 4. Some deep learning models' accuracy

Model	Average accuracy	Note
VGG19	90%	
ResNet10V2	93,8%	
Resnet152V2	94,2%	
MobileNet	89.5%	Model is small-sized and less complex
DenseNet201	93,6%	
InceptionResNetV2	95,3%	Trained models have higher accuracy when the epoch is smaller

Compared to these models, AiCOVID has a higher accuracy rate of 95,765%.

Study limitations

AiCOVID has shown its advantages in detecting COVID-19 with a low risk of transmission and can be used to filter patients with lung injuries at hospitals. However, it cannot be used in cases where the patients have just been infected since the lung's radiographic appearance doesn't show any problem.

Moreover, most publicly available COVID-19 samples are in basic visual formats such as JPEG and PNG, which may cause bias in the prediction due to loss of informational value.

Lacking homogeneousness is also another issue since our dataset samples are collected from various resources. Images should be preprocessed to ensure that the CXR images are free from misleading components, such as alternative texts or numbers. [7]

V. CONCLUSIONS AND FUTURE WORKS

Conclusions

Using deep learning, image multi-classification algorithms, and convolutional neural networks, we built the model called AiCOVID with high accuracy and stability. Moreover, our research also concentrates on post COVID-19 problems and using X-ray images to detect other lung diseases with similar radiographic appearances.

Future works

In the future, our model is predicted to be improved and resolved current issues, contributing to a practically planning scenario for COVID-19:

Figure 1. COVID-19 practically planning scenario

Besides that, we can advance the model in suggesting clinical protocols for COVID-19 patients based on their lungs' radiography appearances and detecting COVID-19 variants.

VI. Acknowledgment

First of all, I would like to give special thanks to MD. Tran Nguyen Thi Tuyen, Radiologist, lecturer at Pham Ngoc Thach Medical University for your precious consulting in medical issues and research methods. Last but not least, I would like to extend my sincere thanks to Ms. Thinh Tran - Computational Biology PhD candidate at Memorial Sloan Kettering for your helpful advises and support when I first started this project.

References

1. Aditya Saxena and Shamsheer Pal Singh, "A Deep Learning Approach for the Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks" (arXiv, January 24, 2022), http://arxiv.org/abs/2201.09952
2. "ACR Recommendations for the Use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection," March 22, 2020, https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection
3. Roger Yat-Nork Chung and Minnie Ming Li, "Anti-Chinese Sentiment during the 2019-NCoV Outbreak," *Lancet (London, England)* 395, no. 10225 (February 29, 2020): 686–87, https://doi.org/10.1016/S0140-6736(20)30358-5
4. Danielle Toussie et al., "Clinical and Chest Radiography Features Determine Patient Outcomes in Young and Middle-Aged Adults with COVID-19," *Radiology* 297, no. 1 (October 2020): E197–206, https://doi.org/10.1148/radiol.2020201754
5. Rikiya Yamashita et al., "Convolutional Neural Networks: An Overview and Application in Radiology," *Insights into Imaging* 9, no. 4 (August 2018): 611–29, https://doi.org/10.1007/s13244-018-0639-9
6. Michael Chung et al., "CT Imaging Features of 2019 Novel Coronavirus (2019-NCoV)," *Radiology* 295, no. 1 (April 2020): 202–7, https://doi.org/10.1148/radiol.2020200230
7. Hanan S. Alghamdi et al., "Deep Learning Approaches for Detecting COVID-19 From Chest X-Ray Images: A Survey," *IEEE Access: Practical Innovations, Open Solutions* 9 (2021): 20235–54, https://doi.org/10.1109/ACCESS.2021.3054484
8. Andre Esteva et al., "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks," *Nature* 542, no. 7639 (February 2, 2017): 115–18, https://doi.org/10.1038/nature21056
9. Babak Ehteshami Bejnordi et al., "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer," *JAMA* 318, no. 22 (December 12, 2017): 2199–2210, https://doi.org/10.1001/jama.2017.14585
10. Jian-Long He et al., "Diagnostic Performance between CT and Initial Real-Time RT-PCR for Clinically Suspected 2019 Coronavirus Disease (COVID-19) Patients Outside Wuhan, China," *Respiratory Medicine* 168 (July 2020): 105980, https://doi.org/10.1016/j.rmed.2020.105980
11. Fengxiang Song et al., "Emerging 2019 Novel Coronavirus (2019-NCoV) Pneumonia," *Radiology* 297, no. 3 (December 2020): E346, https://doi.org/10.1148/radiol.2020209021
12. Ho Yuen Frank Wong et al., "Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19," *Radiology* 296, no. 2 (August 2020): E72–78, https://doi.org/10.1148/radiol.2020201160
13. Fouzia Altaf et al., "Going Deep in Medical Image Analysis: Concepts, Methods, Challenges and Future Directions" (arXiv, February 14, 2019), http://arxiv.org/abs/1902.05655
14. Md Mamunur Rahaman et al., "Identification of COVID-19 Samples from Chest X-Ray Images Using Deep Learning: A Comparison of Transfer Learning Approaches," *Journal of X-Ray Science and Technology* 28, no. 5 (2020): 821–39, https://doi.org/10.3233/XST-200715
15. Tomás Franquet, "Imaging of Pulmonary Viral Pneumonia," *Radiology* 260, no. 1 (July 2011): 18–39, https://doi.org/10.1148/radiol.11092149
16. Ming-Yen Ng et al., "Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review," *Radiology. Cardiothoracic Imaging* 2, no. 1 (February 2020): e200034, https://doi.org/10.1148/ryct.2020200034
17. Harrison X. Bai et al., "Performance of Radiologists in Differentiating COVID-19 from Non-COVID-19 Viral Pneumonia at Chest CT," *Radiology* 296, no. 2 (August 2020): E46–54, https://doi.org/10.1148/radiol.2020200823
18. "Pneumonia of Unknown Cause – China," January 5, 2020, https://www.who.int/emergencies/disease-outbreak-news/item/2020-DON229
19. Adam Jacobi et al., "Portable Chest X-Ray in Coronavirus Disease-19 (COVID-19): A Pictorial Review," *Clinical Imaging* 64 (August 2020): 35–42, https://doi.org/10.1016/j.clinimag.2020.04.001
20. Heshui Shi et al., "Radiological Findings from 81 Patients with COVID-19 Pneumonia in Wuhan, China: A Descriptive Study," *The Lancet. Infectious Diseases* 20, no. 4 (April 2020): 425–34, https://doi.org/10.1016/S1473-3099(20)30086-4
21. Yicheng Fang et al., "Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR," *Radiology* 296, no. 2 (August 2020): E115–17, https://doi.org/10.1148/radiol.2020200432
22. Vincent C. C. Cheng et al., "Severe Acute Respiratory Syndrome Coronavirus as an Agent of Emerging and Reemerging Infection," *Clinical Microbiology Reviews* 20, no. 4 (October 2007): 660–94, https://doi.org/10.1128/CMR.00023-07

23. Pengfei Sun et al., "Understanding of COVID-19 Based on Current Evidence," *Journal of Medical Virology* 92, no. 6 (June 2020): 548–51, https://doi.org/10.1002/jmv.25722
24. "Weekly Operational Update on COVID-19 - 31 July 2020," August 1, 2020, https://www.who.int/publications/m/item/weekly-update-on-covid-19---31-july-2020
25. "WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020," March 11, 2020, https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

APPENDIX 1: List of abbreviations and jargons

- AI: Artificial Intelligence
- X-rays are a type of radiation which are used in medical fields. X-rays create images on films by passing through tissues, organs, bones, etc.
- CT (Computer tomography) is a technology using X-rays scanning through an area of your body and then processed by computers to create 2D or 3D images.
- MRI (Magnetic resonance imaging) is a medical examination performed using magnetic resonance imaging.
- False positive is when the patient is negative but the test result is positive.
- Teachable Machine: a web-based tool to create Machine learning by Google

APPENDIX 2: Data sets sources

- Nodules in Chest X-rays (JSRT): This is the data set made by Japanese Society of Radiological Technology. This data set consists of qualitive X-ray images of lung diseases, which are collected and classified by 20 radiologists. For more details, please visit: Lung nodule data sets descriptions
- COVID-19 Radiography Database: This is a database compiled by researchers from Qatar University, Doha, Qatar and Dhaka University, Bangladesh with some collaborators from Pakistan, Malaysia and medical doctors. This database won the competition "COVID-19 Dataset Award" by Kaggle Community.
- COVID-19 & Pneumonia: This data set is compiled by researchers from GitHub and Kaggle
- COVID-19 X-ray Images
- COVID-19 Chest X-ray Image Repository
- COVID-19 Patients Lung X-ray 1000 Images
- COVID-19 Chest X-ray
- COVID-19 X-ray Image Data set with Huge Sample
- COVID-19 X-ray Data sets
- COVID-19 Chest X-Ray Image Repository