The Importance of Health Informatics in Public Health during a Pandemic
J. Mantas et al. (Eds.)

doi:10.3233/SHTI200481

Setting up an Easy-to-Use Machine Learning Pipeline for Medical Decision Support: A Case Study for COVID-19 Diagnosis Based on Deep Learning with CT Scans

Aikaterini SAKAGIANNI [a], Georgios FERETZAKIS [b,c,1], Dimitris KALLES [b], Christina KOUFOPOULOU [d] and Vasileios KALDIS [e]
a Sismanogleio General Hospital, Intensive Care Unit, Marousi, Greece
b School of Science and Technology, Hellenic Open University, Patras, Greece
c Sismanogleio General Hospital, Department of Quality Control, Research and Continuing Education, Marousi, Greece
d Agia Sofia Children's Hospital, Athens, Greece
e Sismanogleio General Hospital, Emergency Department, Marousi, Greece

Abstract. Coronavirus disease (COVID-19) constitutes an ongoing global health problem with significant morbidity and mortality. It usually presents characteristic findings on a chest CT scan, which may lead to early detection of the disease. A timely and accurate diagnosis of COVID-19 is the cornerstone for the prompt management of the patients. The aim of the present study was to evaluate the performance of an automated machine learning algorithm in the diagnosis of Covid-19 pneumonia using chest CT scans. Diagnostic performance was assessed by the area under the receiver operating characteristic curve (AUC), sensitivity, and positive predictive value. The method's average precision was 0.932. We suggest that auto-ML platforms help users with limited ML expertise train image recognition models by only uploading the examined dataset and performing some basic settings. Such methods could deliver significant potential benefits for patients in the future by allowing for earlier disease detection and care.

Keywords. Artificial intelligence, automated machine learning, AutoML Vision, COVID-19, chest CT scan, image classification, coronavirus disease

1. Introduction and background

Coronavirus disease 2019 (COVID-19) constitutes an ongoing global health problem with significant morbidity and mortality that has emerged in China in December 2019. Due to its rapid dissemination, the World Health Organization (WHO) declared the novel coronavirus outbreak as a pandemic on March 11, 2020. The clinical course of the disease varies and, in severe cases, may lead to acute respiratory distress syndrome

[1] Corresponding Author, Georgios Feretzakis, School of Science and Technology, Hellenic Open University, Patras, Greece; E-mail: georgios.feretzakis@ac.eap.gr.

and/or multiple organ failure. Timely and accurate diagnosis of COVID-19 is of utmost importance for the prompt treatment of the patients and their isolation. The diagnosis is confirmed by reverse-transcription polymerase chain reaction (RT-PCR). According to several reports, the sensitivity of RT-PCR might not be high enough to rapidly recognize infected patients. Computed tomography (CT) of the chest can detect quite early after the onset of symptoms, certain features of lung infection associated with COVID-19. CT could, therefore, serve as a useful tool for the early detection of COVID-19 lung involvement. The National Health Commission of the People's Republic of China has initially proposed screening, based only on clinical and chest CT findings [1].

Applications of artificial intelligence (AI) are rapidly entering health care and are being used for a wide range of health care and research purposes, including, for example, disease detection [2], empirical therapy selection [3], and drug discovery [4], among others. The complexity and the growing volumes of data in health care suggest that AI techniques will increasingly be applied in almost every medical field in the upcoming years. Recent studies have demonstrated that AI using deep learning models may prove extremely helpful in the medical imaging domain due to its high capability in identifying specific disease patterns.

Automated deep learning may provide an opportunity to considerably expand the use of AI in health care. This will encourage medical professionals who are not AI experts to experiment with AI algorithms, allowing them to explore a much wider range of applications [5]. Such methods could deliver significant potential benefits for patients in the future by allowing for earlier disease detection and care. The aim of the present study was to evaluate the performance of an automated machine learning algorithm in the diagnosis of COVID-19 pneumonia using chest CT scans.

2. Methods and Materials

We analyzed a dataset that contains 349 COVID-19 and 397 non-COVID-19 CT scan images, which were captured by patients with any lung diseases, not associated with COVID-19. These images were gathered from COVID-19 articles on medRxiv and bioRxiv. The procedure of the collection of these CTs images is described in detail in [6, 7], and the dataset is available for research purposes [8]. For our experiment, we chose to utilize the Google AutoML Cloud Vision for model design, training, validation, and testing. Neural Architecture Search (NAS) [9] is the underlying framework that supports this platform to find the best way to train deep learning models. AutoML platforms can assist users with limited ML expertise in developing image recognition models by only uploading the examined dataset and performing some basic configuration settings. The number of CTs images that were used in every stage of the model's creation procedure are shown in the following table (Table 1).

Table 1. Number of images that have been assigned to the Train, Validation and Test sets

Label	Train	Validation	Test	Total
COVID-19	279	34	36	349
Non-COVID-19	317	39	41	397
Total	596	73	77	746

3. Results

The AutoML Cloud Vision performance metrics are recall, precision, and average precision. The average precision measures how well the deep learning model performs across all score thresholds and is the area under the precision-recall curve. The closer this score is to 1.0, the better the model performs on the test set. In our case, the average precision's value is 0.932. The recall (sensitivity) is 0.8831, and precision (positive predictive value) is 0.8831 for a given score threshold.

The score threshold, in our example, was set 0.5, allows us to explore how this value affects the precision and recall of our model, and the corresponding tradeoff curve (Figure 1). This can help us find a good balance between false positives and false negatives. In our case, by having set the score threshold 0.5, the precision and recall both have the same value (=0.8831), which leads to an F1 measure equals to 0.8831.

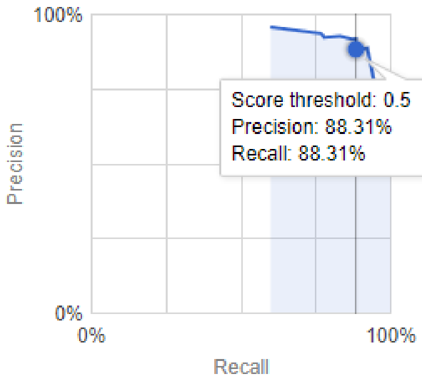

Figure 1. Precision-Recall tradeoff curve

Additionally, the confusion matrix for the deep learning model is shown below (Table 2), which represents the cross-tabulating ground truth labels versus the labels predicted by the model.

Table 2. Confusion Matrix of 77 test CTs images

	Predicted Label	
True Label	COVID-19	Non-COVID-19
COVID-19	31	5
Non-COVID-19	4	37

4. Discussion

Current literature suggests that lung imaging may play a pivotal role in screening and in the rapid diagnosis of COVID-19 disease, in symptomatic patients, pending confirmation by molecular tests [10-12]. In this study, we evaluated the performance of an AutoML Cloud Vision model with promising results in terms of accuracy in detecting patients infected with COVID-19 using CT lung scans. Our results obtained by using AutoML are comparable with the performance ($F1 = 0.85$ and $AUC = 0.94$) of classic deep learning models developed using traditional non-autoML techniques [7] applied to the same dataset. Despite the high positive predictive value of lung CT findings in coronavirus disease, certain limitations exist. COVID-19 pneumonia shares common imaging characteristics with other viral pneumonias or non-infectious lung diseases. Moreover, in the early stages of the disease or in asymptomatic patients, a CT scan may be negative. Therefore, a multidisciplinary approach to diagnosis is essential.

5. Conclusions

There are several constraints that limit the development and widespread adoption of AI methods in clinical practice, with the major concern being the lack of open datasets to train AI models. Collaborative projects which gather and share big data may contribute to the fight against this new pandemic. AutoML platforms may help users with limited ML expertise train image recognition models by only uploading the examined dataset and performing some basic settings. This could prove a substantial tool in the quest of the medical profession to understand and treat the disease, eventually delivering significant potential benefits for patients in the future.

References

[1] Yijiu X, China's Hubei reports jump in new cases of COVID-19 after diagnosis criteria revision. National Health Commission of the People's Republic of China website, www.en.nhc.gov/cn/2020-02/13/c_76515.htm (accessed on 23 April 2020).

[2] Sharma P, Choudhary K, Gupta K, Chawla R, Gupta D, Sharma A, Artificial Plant Optimization Algorithm To Detect Heart Rate & Presence Of Heart Disease Using Machine Learning, Artificial Intelligence In Medicine, 102(101752) (2019).

[3] Feretzakis G, Loupelis E, Sakagianni A, Kalles D, Martsoukou M, Lada M, Skarmoutsou N, Christopoulos C, Valakis K, Velentza A, Petropoulou S, Michelidou S, Alexiou K, Using Machine Learning Techniques to Aid Empirical Antibiotic Therapy Decisions in the Intensive Care Unit of a General Hospital in Greece, Antibiotics 9(2) (2020), 50.

[4] Mak KK, Pichika MR, Artificial Intelligence In Drug Development: Present Status And Future Prospects, Drug Discovery Today, 24(3) (2019), 773-780.

[5] Faes L et al., Automated Deep Learning Design for Medical Image Classification by health-care professionals with no coding experience: a feasibility study, The Lancet Digital Health, 1(5) (2019), 232-242.

[6] Zhao J, Zhang Y, He X, Xie P, COVID-CT-Dataset: A CT Scan Dataset about COVID-19, ArXiv:2003.13865 (2020).

[7] He X, Yang X, Zhang S, Zhao J, Zhang Y, Xing E, Xie P, Sample-Efficient Deep Learning for COVID-19 Diagnosis Based on CT Scans, MedRxiv 2020.04.13.2006394 (2020).

[8] COVID-CT-Dataset: A CT Scan Dataset about COVID-19 https://github.com/UCSD-AI4H/COVID-CT (accessed on 21April 2020).

[9] Elsken T, Metzen JH, Hutter F, Neural Architecture Search: A Survey, ArXiv:1808.05377, Apr. 2019. arXiv.org, http://arxiv.org/abs/1808.05377.

[10] Pan F, Ye T, Sun P, Gui S, Liang B, Li L, Zheng D, Wang J, Hesketh RL, Yang L, Zheng C, Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia, Radiology 0 (2020).

[11] Zu ZY, Jiang MD, Xu PP, Chen W, Ni QQ, Lu GM, Zhang LJ, Coronavirus Disease 2019 (COVID-19): A Perspective from China, Radiology (2020).

[12] Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q. Sun Z, Xia L, Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases, Radiology (2020).