# Multi-classification machine learning for diagnosing COVID-19 in X-ray

**Mayada Jabbar Kelain**

AL -Mustansiriyah University College of Tourism Sciences
Mayadajabbar@uomustansiriyah.edu.iq

ABSTRACT

The dangerous COVID-19 virus is a threat to all human beings around the world. Effective identification of COVID-19 using advanced machine learning methods is a timely need. Although many complex methods have been proposed in the recent past, they still struggle to achieve the expected performance in classifying and identifying COVID-19 patients using chest X-rays. In addition, most of them are involved in the complex pre-treatment task, which is often difficult for a virologist. Meanwhile, deep networks are comprehensive and have shown promising results in image recognition tasks over the past decade. In this work, chest x-ray images were used after processing the images using filters, as well as determining the infection with the virus and its classification by the SVM algorithm, as the algorithm gave good and effective results in knowing the person infected with corona or not.

**Keywords:**

## 1. Introduction

A new virus has been shown to spread from person to person. Through the respiratory tract, the virus enters the lungs where it spreads through the entire lung until it reaches the alveoli. The immune system will overreact once it binds to a receptor because it is unable to identify that pathogen. As a result, a large number of immune cells are pushed into the lung, doing more harm than good. It is critical because this virus may affect the immune system. It can affect every part of the respiratory system, including the alveoli, which cover 99% of the surface area of the lung. The coronavirus interferes with the mechanism by which the alveoli remove carbon dioxide and transport oxygen into the circulatory system. In the worst case, the walls of the alveoli begin to collapse. Thus, the cells of the lungs and fluid begin to fill the lungs.

When this happens, the patient cannot excrete enough gas, which prevents gas exchange, nor adequate oxygen or carbon dioxide capture. Breathing became increasingly difficult [2].

In the event of a diagnosis of COVID-19, infected individuals must be separated very quickly to reduce the spread of the virus. Reverse transcription-polymerase chain reaction (RTPCR) is an essential method for diagnosing COVID-19, which involves detecting viral RNA in nasopharyngeal or sputum swabs. There are several limitations to the reverse transcription-polymerase chain reaction (RTPCR) tests, including the long time they take. A low positive detection rate has been reported according to the World Health Organization (WHO). As a result of these limitations, the infectious nature of the disease may lead to viral infection in healthy individuals. As an alternative method, chest CT imaging has been used to diagnose COVID-19 infection because radiographic specimens on chest CT images showed higher sensitivity and specificity in detecting COVID19 and are cheap and easy to use. Despite the impact of COVID-19 in all facilities of life, including the deterioration of the economy and trade, and its impact on education in schools and universities around the world.

Implementation of an automatic detection system as an alternative diagnostic option can prevent the spread of COVID-19 [3].

The field of artificial intelligence has expanded rapidly with the use of deep neural networks to solve various problems such as object detection, speech recognition, and image classification. Specifically, convolutional neural networks (CNNS) performed well at image classification. Many works have been done to detect and diagnose COVID-19 with AI including machine learning and deep learning as it is the best technology for early detection and disease diagnosis. Deep learning for reliable diagnostic accuracy in health imaging for the automated detection of lung diseases has been heralded as one of the most important artificial intelligence (AI) techniques [4]

## 2. Literature Review

1- Loey, Gunasekaran, and Nour, 2020 [6] used five different deep CNN models (ResNet50, GoogleNet, VGGNet19, VGGNet16, and for COVID-19 pneumonia use AlexNet. The research report shows how to use data augmentation techniques with model networks Conditional generative adversarial based on deep learning (CGAN) for diagnosing COVID-19 from chest CT scans The primary problem is the limited size of the COVID-19 CT database Best results from this study assuming classification specificity of the ResNet50 model Effectiveness 80.56% Sensitivity, 85.87% Accuracy, 90.42%.

2- Jim et al., 2020 [7] A novel sequential convolutional neural network identified using a computed tomography (CT) scan image analysis model, COVID-19. This dataset contains 349 tomography images from 216 individuals with the condition. The model is able to locate disease with a sensitivity of 94%, F1 score 94%, accuracy 94%, accuracy 92% and accuracy, 93%.

3- Mishra et al., 2020 [8] investigated different CNN-based methods for detecting COVID-19 on x-ray images. In the decision fusion based architecture, several models (DenseNet201, DenseNet121, InceptionV3, ResNet50, and VGG16) are combined by majority voting to produce an overall prediction. The prediction accuracy of the decision fusion-based strategy as a whole is 86%.

## 3. Objectives of the study:

The primary objective of this paper is to develop a mobile diagnostic application that can identify COVID-19 patients by looking at images of a patient's lungs obtained from a CT scan. A COVID-19 patient can be distinguished from other patients using image processing techniques and artificial intelligence algorithms. The proposed study will reduce the time required for diagnosis, which is often carried out by specialists in communicable diseases and diseases of the respiratory system

## 4. Extraction Features

The feature extraction process was used to classify pneumonia as well as covid-19, or the chest cage is normal. In this paper, feature extraction in the proposed structure of the machine learning technique consisted of three consecutive Conv blocks where the most accurate features were selected by which the disease could be classified.

## 5. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised machine learning algorithm that can be used for classification and regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data element as a point in n-dimensional space (where n is the number of features you have) with the value of each feature being a given coordinate value. Next, we classify by finding the super level that distinguishes the two classes well and this algorithm works as follows:

- Determine the correct hyperplane (scenario -1): Here we have three hyperplanes (A, B, and C). Now, select the appropriate super level for star and circle classification.
- Determine the correct hyperplane (scenario -2): Here, we have three hyperplanes (A, B, C) and they all separate the classes well.
- Determining the correct super level (Scenario 3): Hint: Use the rules as

discussed in the previous section to select the correct super level.

- Can we classify two classes (Scenario-4)?: Below, I cannot separate the two classes using a straight line, where one of the stars is located in the territory of another class (circle) as outer.
- As I mentioned earlier, one star at the other end is like a star category. The SVM algorithm has a feature to ignore outliers and find the hyper level that has the maximum margin. Hence, we can say, SVM classification is robust for outliers.
- Find the class separation super level (scenario 5): In the scenario below, we can't have a linear super level between the two classes, so how does SVM classify these two classes? So far, we have only looked at the hyper linear plane ,SVM can solve this problem. Easily It solves this problem by introducing an additional feature.

## 6. X-Ray image datasets

The proposal is based on X-ray images collected from 216 COVID-19 patients posted to (Zhao, et al., 2020) from Dr. Joseph Cohen's open source GitHub repository. The dataset consisted of 746 chest and lung images with (349 COVID-19 images) and (397 non-COVID-19 images). The data set contains details and data for people with covid-19 diseases. The data sets are divided into two samples: the training data make up 80% of the total data set (597 CT scans); while data test makes up 20% (149 CT scans)

## 7. Result

Table (1)

| Type Covid-19 | Number of image | Accuracy ratio | Specificity ratio | Precision ratio | Time execution |
|---|---|---|---|---|---|
| Positive | 50 | (98.33 %) | 96.71 %) | 95.24 %) | 23 Second |
| Negative | 50 | 97.92 %) | 95.43 %) | 93.81 %) | 23 Second |

For testing data, F-measure, Accuracy, Sensitivity, Precision and time have all been measured. As shown in Table (1) the values gradually rise with each epoch. The Testing-Accuracy had reached 98.33%. The accuracy improves with It gets better when using algorithms that help the core algorithm. The following shows the final results of precision 95.24%, specificity 96.71% and time 23 seconds for when the machine learning structure is trained

## 8. Conclusion

COVID-19 is a virus that has spread internationally. This work explores a novel mobile application for COVID-19 detection using a machine learning solution to assist medical professionals. X-rays in X-ray examinations provide reliable, effective and accurate solutions in medical diagnosis. The proposed system aims to detect COVID-19 in chest x-ray images. One problem with this system is the limited data set.

1- The proposed system supports the development of a computer-based system to assist physicians specializing in diagnosis.

2- Because there is so much inflammation in the numerous radiographs that pneumonia virus disease and other lung diseases share, the differential diagnosis is often difficult; Computed tomography, the gold standard in detecting lung diseases caused by the Corona virus, is difficult, especially for people who smoke.

## References

1. Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., & Zhang, L.(2020). Epidemiological and clinical

characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet (London, England), 395(10223), 507–513. https://doi.org/10.1016/S0140-6736(20)30211-7

2. Jenny, O., Laura, K., "How the Coronavirus Attacks Your Lungs". May 5, 2020. Date of retrieval 12-1-2021. https://www.kqed.org/science/1963200/ how -Covid-19 -attacks-yours's-lungs.
3. Purohit, K., Kesarwani, A., Kisku, D. R., & Dalui, M. (2020). "Covid19 detection on chest x-ray and ct scan images using multi-image augmented deep learning model". BioRxiv.
4. Suzuki, K. (2017). "Overview of deep learning in medical imaging".
5. Radiological physics and technology, 10(3), 257-273.
6. Swapnarekha, H., Behera, H. S., Nayak, J., & Naik, B. (2020). Role of intelligent computing in COVID-19 prognosis: A state-of-the-art review. Chaos, solitons, and fractals, 138, 109947.
7. Amyar, A., Modzelewski, R., Li, H., & Ruan, S. (2020). Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation. Computers in biology and medicine, 126, 104037.
8. Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., ... & Li, L. (2020). A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering, 6(10), 1122-1129.
9. Anwar, T., & Zakir, S. (2020, November). Deep learning based diagnosis of COVID-19 using chest CT-scan images. In 2020 IEEE 23rd International Multitopic Conference (INMIC) (pp. 1-5). IEEE.
10. Yang, X., He, X., Zhao, J., Zhang, Y., Zhang, S., & Xie, P. (2020). COVID-CTdataset: a CT scan dataset about COVID-19. arXiv preprint arXiv:2003.13865.
11. Yang, S., Jiang, L., Cao, Z., Wang, L., Cao, J., Feng, R., Zhang, Z., Xue, X., Shi, Y., & Shan, F. (2020). Deep learning for detecting corona virus disease 2019 (COVID-19) on high-resolution computed tomography: a pilot study. Annals of translational medicine, 8(7), 450. https://doi.org/10.21037/atm.2020.03.132
12. Hu, R., Ruan, G., Xiang, S., Huang, M., Liang, Q., & Li, J. (2020). Automated diagnosis of covid-19 using deep learning and data augmentation on chest ct. medRxiv.
13. Wu, X., Hui, H., Niu, M., Li, L., Wang, L., He, B., ... & Zha, Y. (2020). Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: a multicentre study. European Journal of Radiology, Vol. 128,109041.