

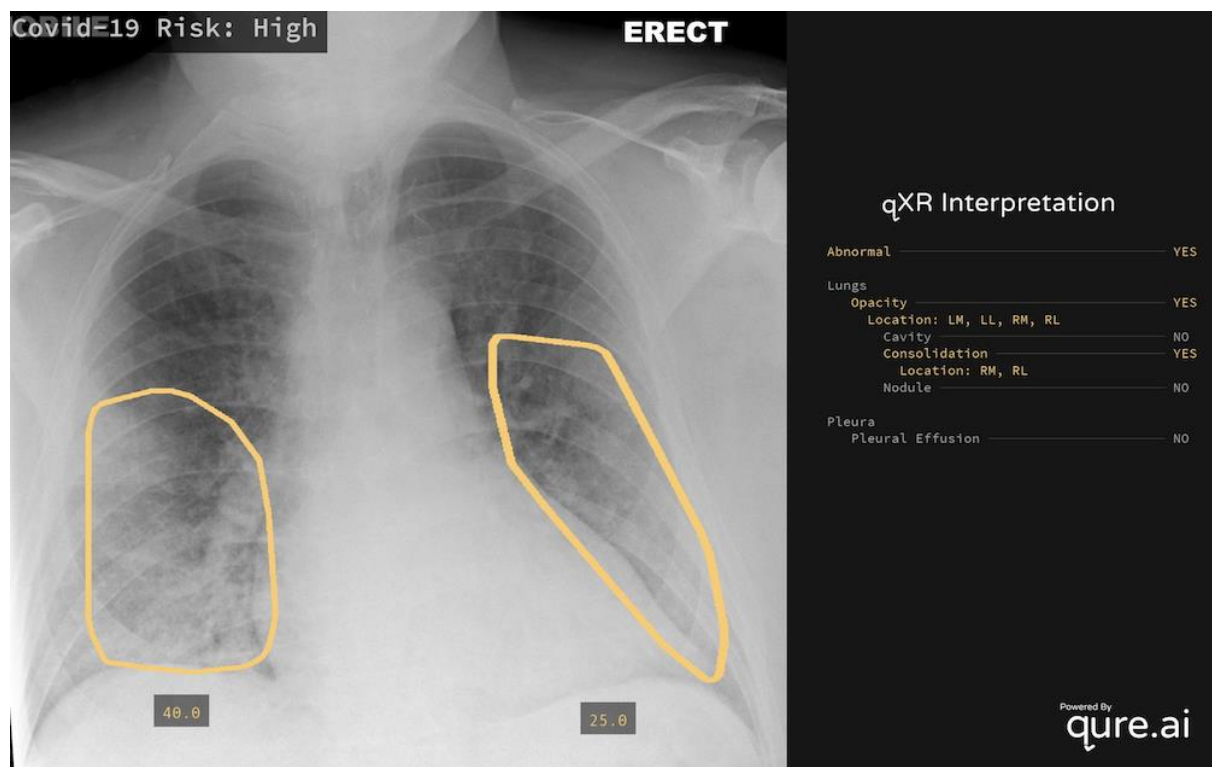
Help health systems manage the pandemic.

Background

When the COVID-19 pandemic hit India in March 2020, it sent the healthcare system into a state of shock. The biggest need of the hour was to enable hospitals deal with the heavy patient inflow. The problem faced by the hospitals was three-fold – 1. Triage patients quickly for quarantine to prevent further spread despite the severe shortage of RT-PCR tests. 2. If someone is suspected to be positive, decide whether to admit for intensive care or send away for quarantine. 3. Decide when to discharge a patient from intensive care to free up space for those who need it the most.

Role of Chest X-rays

Chest X-rays (CXRs) are the most common initial diagnostic tool for most thoracic symptoms. More than a billion chest X-rays are done every year across the world. In terms of assessing the thorax, the Chest X-ray is very cost-effective and is widely available. Initial studies at the beginning of the pandemic indicated that ground glass opacities and consolidations which are peripheral and bilateral in nature are attributed to be the most common findings with respect to COVID-19 related infections on CXRs. Chest X-rays can also help assess the severity of the disease and monitor progression and response to treatment. In view of these, Chest X-rays emerged as the frontline screening, diagnostic and monitoring tool for COVID-19 infection in conjunction with clinical history and blood markers.

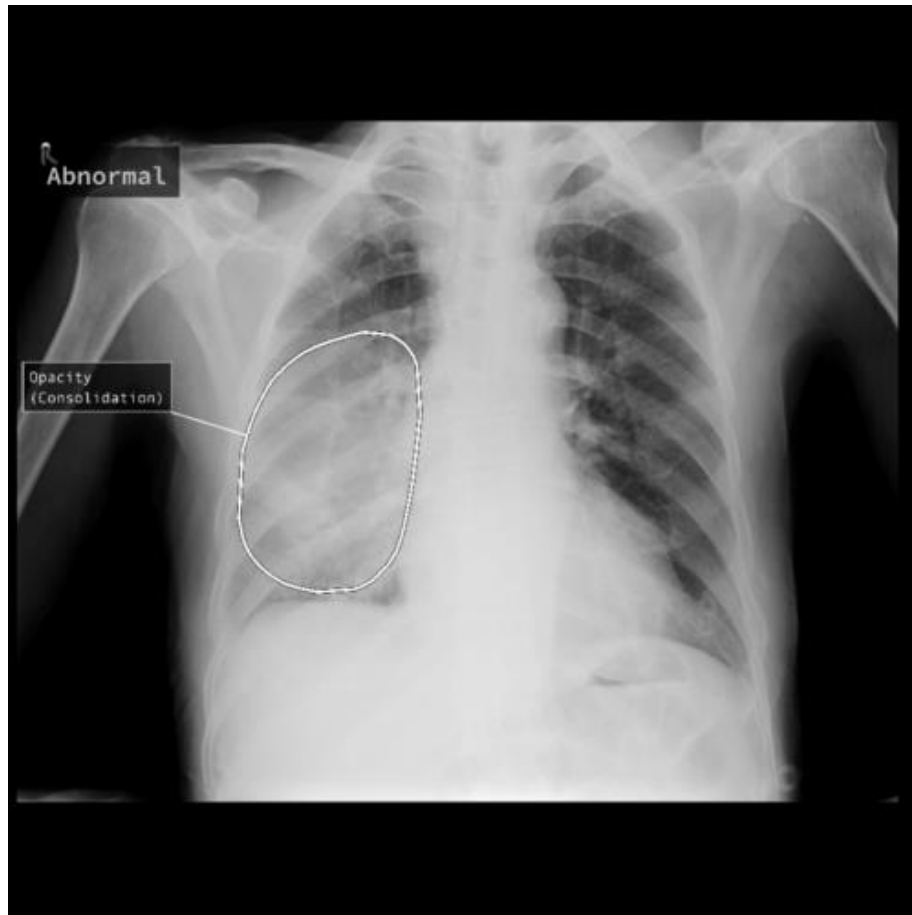


A Chest X-ray of a patient infected with COVID-19.

Problem statement

To enable the use of Chest X-rays at scale, particularly in resource-constrained settings with limited availability of experts, we need an AI algorithm that can detect and quantify findings that are consistent with a COVID-19 infection.

Consolidation is the most important finding when it comes to COVID-19 on Chest X-rays. By definition, A pulmonary consolidation is **a region of normally compressible lung tissue that has filled with liquid instead of air**. The condition is marked by induration (swelling or hardening of normally soft tissue) of a normally aerated lung.



A Chest X-ray with pulmonary consolidation.

For the purposes of this assignment, you will attempt to create an AI algorithm that can detect consolidation from a chest X-ray. You are provided with a dataset of 20000 Chest X-rays in a Png format. Each X-ray has a binary label associated with it indicating 1 if a consolidation is present in it. Additionally, some positive samples(images with consolidation) have information on the side in which consolidation is present. Ground truth data for 20% of the data(4000 samples) is withheld to evaluate your submission. Your goal is to create an algorithm that gives the following outputs.

1. A confidence score indicating the probability that a given Chest X-ray has consolidation.

2. (Bonus) An output indicating the side in which consolidation is present(left, right, both).

Data:

1. You are provided with a zip file `cxr_plaksha_assignment_qure.zip`. Each image [0...20000].png is a Chest X-ray that is resized to a standard size of 1440 x 1440 pixels. Download the zip file using this link - [cxrs.zip](#)
2. The ground truth data is in the form of a csv file with the following header
filename,consolidation,consolidation-left,consolidation-right
4516,1,1,1
18224,0,0,0
2458,1,-100,-100

The second row indicates that the image 4516.png has consolidation, additionally, it also indicates that there is consolidation in both left and the right lungs. The last row indicates that there is consolidation in the image but information about the laterality(side information) is unknown.

The csv has ground truth data for 16000 cases. The other 4000 are set aside as test. Download the csv file using this link - [train_gt.csv](#)