COVID-Xray Dataset

The COVID-X-ray dataset consists of chest X-rays from two datasets: Covid-19 Radiograhpy Dataset which contains a total of 219 Covid-19 Positive images and 1341 Covid-19 Negative test images are separated in two labelled folders. Another dataset is the Covid-Chestxray-Dataset, which was recently published and is comprised of a set of images from published sources regarding COVID-19 topics collected by https://github.com/ ieee8023/covid-chestxray-dataset, Cohen et al. (2020). This dataset is compiled of a blend of CT images along with chest X-rays.

***(Metadata) visualization figure shows that CT Scan was 14%).***

This pool of data is unceasingly updated, and meta-data such as age, findings, sex, survival and hospitalization status are also recorded. Part of the COVID-19 images used for this project contains the images that derived from this dataset. We have only kept the PA view images for COVID-19 prediction*.*

The resolution should be mentioned regarding this dataset, as it varies widely. The lower-resolution COVID-19 images are below 400 x 400, while some of the higher ones are exceeding 1900 x 1400. The models produce accurate results regardless of this disparity in image quality. Although exceptional image quality should be strived for, it is a focus in the machine learning realm to create systems that are able to operate well regardless of differing image qualities. The dynamic ranges also differ as they are images collected from various providers. These images are indeed normalized to decrease model confusion.