**Data Description:**

The data was kindly provided by [medicalsegmentation.com](http://medicalsegmentation.com/covid19/). They are two radiologists from Oslo, who've done plenty of work scraping and segmenting CT images. Also, they have made a [great product](https://www.medseg.ai/), where you can segment CT images in a browser. The magic of their tool is that it has several pretrained Unets that run in offline mode (no user data is going out) and make your work a hundred times faster. An additional contribution of lung masks was made by Johannes Hofmanninger. His approach on [Github](https://github.com/JoHof/lungmask).

**Dataset consists of two parts:**

**Medseg part**

This is a dataset of 100 axial CT images from >40 patients with COVID-19 that were converted from openly accessible JPG images found [HERE](https://www.sirm.org/en/). The conversion process is described in detail in the following blogpost: [Covid-19 radiology — data collection and preparation for Artificial Intelligence](https://medium.com/@hbjenssen/covid-19-radiology-data-collection-and-preparation-for-artificial-intelligence-4ecece97bb5b)

images\_medseg.npy - training images – 100 slices 512x512 size  
masks\_medseg.npy - training masks – 100 masks with 4 channels: (0 - "ground glass", 1 - "consolidations", 2 - "lungs other", 3 - "background" )  
test\_images\_medseg.npy - test images – 10 slices 512x512 size

**Radiopedia part**

Segmented 9 axial volumetric CTs from [Radiopaedia](https://radiopaedia.org/articles/covid-19-3). This dataset includes whole volumes and includes, therefore, both positive and negative slices (373 out of the total of 829 slices have been evaluated by a radiologist as positive and segmented). These volumes are converted and normalized in a similar way as above.

**images\_radiopedia.npy** - training images – 829 slices 512x512 size

**masks\_radiopedia.npy** - training masks – 829 masks with 4 channels: (0 - "ground glass", 1 - "consolidations", 2 - "lungs other", 3 - "background" )

predict 10 masks corresponding to test\_images\_medseg.npy with two classes: 0 - "ground glass", 1 - "consolidations" to validate results .

**Kaggle Api command : kaggle competitions download -c covid-segmentation**