

## Data Collection and Preprocessing Phase

Date	16 March 2024
Team ID	SWTID1720165000
Project Title	CovidVision: Advanced COVID-19 Detection From Lung X-Rays With Deep Learning
Maximum Marks	6 Marks

### Preprocessing Template

The images will be preprocessed by resizing, normalizing, augmenting, denoising, adjusting contrast, detecting edges, converting color space, cropping, batch normalizing, and whitening data. These steps will enhance data quality, promote model generalization, and improve convergence during neural network training, ensuring robust and efficient performance across various computer vision tasks.

Section	Description
Data Overview	The Data contains lung x-rays images and its masks .The dataset has a total of 40,000 images.
Resizing	224 x 224
Normalization	0-1
Data Augmentation	Shearing, rescale, zooming, horizontal_flip.
Denoising	Nil
Edge Detection	Nil

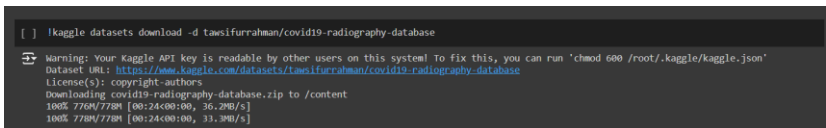
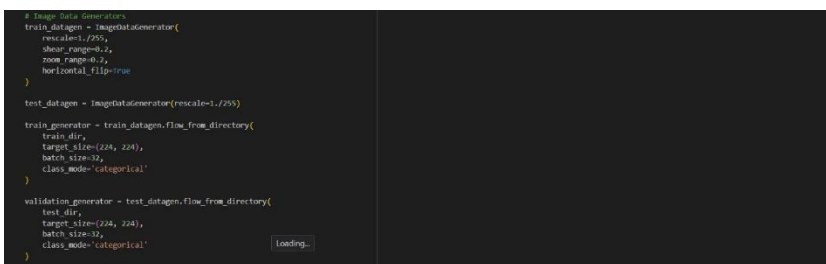
Color Space Conversion	Nil
Image Cropping	224 x 224
Batch Normalization	32
<b>Data Preprocessing Code Screenshots</b>	
Loading Data	 <pre>[ ] kaggle datasets download -d tawsifurrahman/covid19-radiography-database Warning: Your Kaggle API key is readable by other users on this system. To fix this, you can run 'chmod 600 /root/.kaggle/kaggle.json' Dataset URL: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database License(s): copyright-authors Downloading covid19-radiography-database.zip to /content 100% 776M/776M [00:24&lt;00:00, 36.2MB/s] 100% 776M/776M [00:24&lt;00:00, 33.3MB/s]</pre>
Resizing	 <pre># Image Data Generators train_datagen = ImageDataGenerator(     rescale=1./255,     shear_range=0.2,     zoom_range=0.2,     horizontal_flip=True )  test_datagen = ImageDataGenerator(rescale=1./255)  train_generator = train_datagen.flow_from_directory(     train_dir,     target_size=(224, 224),     batch_size=32,     class_mode='categorical' )  validation_generator = test_datagen.flow_from_directory(     test_dir,     target_size=(224, 224),     batch_size=32,     class_mode='categorical' )</pre>
Normalization	Nil
Data Augmentation	Nil
Denoising	Nil
Edge Detection	Nil
Color Space Conversion	Nil

Image Cropping	<pre>train_generator = train_datagen.flow_from_directory(     train_dir,     target_size=(224, 224),     batch_size=32,     class_mode='categorical' )</pre>
Batch Normalization	<pre>train_generator = train_datagen.flow_from_directory(     train_dir,     target_size=(224, 224),     batch_size=32,     class_mode='categorical' )</pre>