Data Collection and Preprocessing Phase

Date	08-05-2024
Team ID	SWTID1720433291
Project Title	CovidVision: Advanced COVID-19 Detection From Lung X-Rays With Deep Learning
Maximum Marks	6 Marks

Preprocessing :

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	Uses lung X-ray images for COVID-19 detection. The dataset consists of images in various formats, including training, validation, and testing sets.
Resizing	Images are resized to a target size of 150x150 pixels and 256x256 for another model
Normalization	Values are normalized using xception preprocessing function.
Data Augmentation	Augmentation techniques include random horizontal flipping, rotation, and zooming to improve model generalization.
Denoising	Not done explicitly. It is done during data augmentation

Edge Detection	Not applied as it learns features through layers
Color Space Conversion	Convert images from one color space to another. Applied within the preprocessing layers itself
Image Cropping	Not applied .Model works with resized images directly
Batch Normalization	Applied within the architecture itself
Data Preprocessing Code Screenshots	
Loading Data	
Data Overview	

```
train_datagen = ImageDataGenerator(preprocessing_function=tf.keras.applications.xception.preprocess_input,zoom_range=0.1,brightness_range=[0.5,1.3],
                                   width_shift_range=0.1,height_shift_range=0.1,validation_split=0.1)
test_datagen=ImageDataGenerator(preprocessing_function=tf.keras.applications.xception.preprocess_input)
```

```
[6]: df_meta=pd.read_csv("Chest_xray_Corona_Metadata.csv",index_col=0)
     df_meta.head()
```

[6]:

	X_ray_image_name	Label	Dataset_type	Label_2_Virus_category	Label_1_Virus_category
0	IM-0128-0001.jpeg	Normal	TRAIN	NaN	NaN
1	IM-0127-0001.jpeg	Normal	TRAIN	NaN	NaN
2	IM-0125-0001.jpeg	Normal	TRAIN	NaN	NaN
3	IM-0122-0001.jpeg	Normal	TRAIN	NaN	NaN
4	IM-0119-0001.jpeg	Normal	TRAIN	NaN	NaN

Resizing Normalization Data Augmentation Denoising Color Space Conversion Batch Normalisation	
Image Cropping Edge Detection	-