Data Collection and Preprocessing Phase

Date	15 March 2024
Team ID	SWTID1720425899
Project Title	Covid Vision: Advanced COVID-19 Detection from Lung X-rays with Deep Learning
Maximum Marks	6 Marks

Preprocessing:

Section	Description
Data Overview	The dataset contains data of chest X-ray images for Covid-19 positive cases along with Normal and Viral Pneumonia images.
Resizing	Resize images to a target size of (256, 256)
Normalization	rescale=1./255: This parameter is used to normalize pixel values of images. In many image datasets, pixel values range from 0 to 255.
Data Augmentation	Applied augmentation techniques such as zooming range = 0.2 , shearing range = 0.2.
Data Preprocessing Code Screenshots	
| Loading Data | ```
!pip install opendatasets
import opendatasets as od
import pandas as pd
od.download("https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database")
``` |

| | ```
!pip install opendatasets
import opendatasets as od
import pandas as pd
od.download("https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database")

Requirement already satisfied: opendatasets in /usr/local/lib/python3.10/dist-packages (0.1.22)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from opendatasets) (4.66.4)
Requirement already satisfied: kaggle in /usr/local/lib/python3.10/dist-packages (from opendatasets) (1.6.14)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from opendatasets) (8.1.7)
Requirement already satisfied: six>=1.10 in /usr/local/lib/python3.10/dist-packages (from kaggle->opendatasets) (1.16.0)
Requirement already satisfied: certifi>=2023.7.22 in /usr/local/lib/python3.10/dist-packages (from kaggle->opendatasets) (2024.6.2)
Requirement already satisfied: python-dateutil in /usr/local/lib/python3.10/dist-packages (from kaggle->opendatasets) (2.8.2)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from kaggle->opendatasets) (2.31.0)
Requirement already satisfied: python-slugify in /usr/local/lib/python3.10/dist-packages (from kaggle->opendatasets) (8.0.4)
Requirement already satisfied: urllib3 in /usr/local/lib/python3.10/dist-packages (from kaggle->opendatasets) (2.0.7)
Requirement already satisfied: bleach in /usr/local/lib/python3.10/dist-packages (from kaggle->opendatasets) (6.1.0)
Requirement already satisfied: webencodings in /usr/local/lib/python3.10/dist-packages (from bleach->kaggle->opendatasets) (0.5.1)
Requirement already satisfied: text-unidecode>=1.3 in /usr/local/lib/python3.10/dist-packages (from python-slugify->kaggle->opendatasets) (1.3)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->kaggle->opendatasets) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->kaggle->opendatasets) (3.7)
Skipping, found downloaded files in "./covid19-radiography-database" (use force=True to force download)
``` |
| Resizing | ```
IMAGE_SIZE = (256, 256)
IMAGE_SHAPE = IMAGE_SIZE + (3,)
``` |
| Normalization | ```
from tensorflow.keras.preprocessing.image import ImageDataGenerator
img_height, img_width= IMAGE_SIZE
batch_size=16
train_datagen = ImageDataGenerator(rescale=1./255,zoom_range=0.2,shear_range=0.2,validation_split=0.3)
``` |
| Data Augmentation | ```
from tensorflow.keras.preprocessing.image import ImageDataGenerator
img_height, img_width= IMAGE_SIZE
batch_size=16
train_datagen = ImageDataGenerator(rescale=1./255,zoom_range=0.2,shear_range=0.2,validation_split=0.3)
``` |