

Build an Image Segmentation Model using Amazon SageMaker

Overview

Amazon Sagemaker is a machine learning service where data scientists can train and deploy their machine learning models. It provides an integrated Jupyter authoring notebook instance to access the data sources for exploration and analysis efficiently. To make the work of data scientists easy, it provides easy deployment of the model in the hosted environment so that they don't have to rely upon software engineers for it.

Sagemaker enables jobs to preprocess, and post-process the data. The user can also do feature engineering, and evaluate their models.

In this project, we will build an image segmentation model in Tensorflow on amazon sagemaker using the UNet model architecture. The project will also be deployed on the sagemaker. You can see the previous project of the series [Build a Text Generator Model using Amazon SageMaker](#)

Aim

- To understand image segmentation
- To train and deploy the image segmentation model on amazon sagemaker

Tech Stack

- Language: Python
- Libraries: tensorflow, pandas, numpy, sklearn, patchify, matplotlib, segmentation_models, boto3

Data Description

The dataset contains images of apples that are rotten. The images will be used to train the model. The rotten part of the apples is annotated.

Approach

- Data Loading
- Data Preprocessing
 - Image Patching
- Model Building and Training

- Segmentation model in tensorflow
- Model Deployment on Amazon Sagemaker

Modular Code Overview

data

|_images

MLPipeline

|_Data_Processing.py

|_Loading_Data.py

|_Model_Training.py

|_Plotting.py

NoteBook

|_Image_Segmentation.ipynb

Sagemaker_Deployment

|_Deployement_ipynb

|_ImageSegmentation_Deployment.ipynb

|_Image_Segmentation_py

|_ImageSegmentation.py

Engine.py

Readme.md

requirements.txt

Once you unzip the image_segmentation.zip file, you can find the following folders within it.

1. Input
 2. MLPipeline
 3. NoteBook
 4. Sagemaker_Deployment
 5. Engine.py
 6. Readme.md
 7. requirements.txt
-
1. The Input folder contains the data that we have for analysis. In our case, it includes images of rotten apples.

2. The Notebook folder contains the jupyter notebook file of the project
3. The ML_pipeline is a folder that contains all the functions put into different python files, which are appropriately named. These python functions are then called inside the Engine.py file
4. Sagemaker_Deployment folder contains the files for deployment
5. The requirements.txt file has all the required libraries with respective versions. Kindly install the file by using the command **pip install -r requirements.txt**
6. **All the instructions for running the code are present in Readme.md file**

Takeaways

1. What is computer vision?
2. What is image segmentation?
3. Introduction to TensorFlow.
4. What is Semantic segmentation and Instance segmentation?
5. Introduction to CNN.
6. What is UNet model architecture?
7. Advantages of UNet architecture.
8. What is sagemaker?
9. How sagemaker work?
10. How to create a notebook instance in sagemaker?
11. How to upload data on the s3 bucket?
12. How to fetch data from the s3 bucket?
13. How to patchify the images for segmentation?
14. Model training in sagemaker
15. Deploying model on sagemaker.