**Text extraction from image using Optical character recognition**

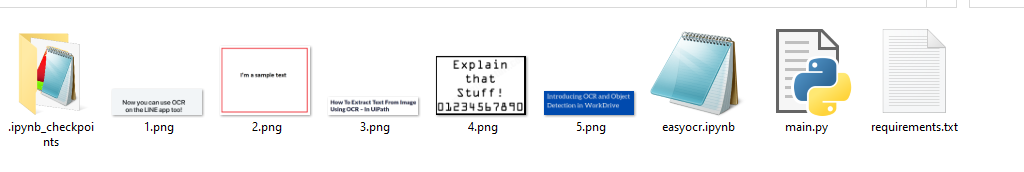
How many of us would have wanted to get information from a YouTube video or pdf or any image. This program helps us achieve this. You just have to give the image and the program will return whatever text is present in the image. Optical Character Recognition helps us achieve it in a very easy way.

**Base Paper:**

### <http://ceur-ws.org/Vol-2870/paper15.pdf>

### Dataset Description:

No dataset used. For testing the accuracy of our OCR reader, I have given some images.



**Algorithm Description:**

EasyOCR:

As the name suggests EasyOCR is the most easy and efficient way to read text from an image. OCR refers to Optical Character Recognition which helps you to read text from an image directly. We might have felt the need to copy the text from youtube but there is no such option on it. Using OCR, we can get the text by passing the screenshot and retrieve the text. EasyOCR works on more than 50 languages including Hindi, Russian and many more.

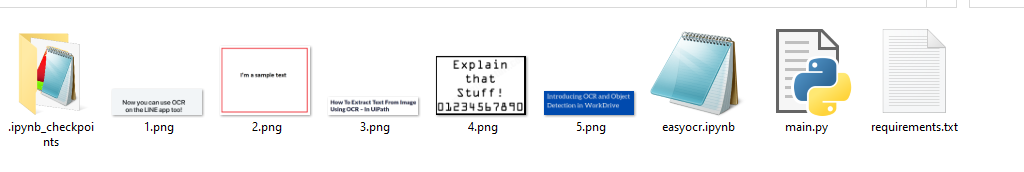
To read more about EasyOCR, you can refer to this link https://www.analyticsvidhya.com/blog/2021/06/text-detection-from-images-using-easyocr-hands-on-guide/

# How to Execute?

So, before execution we have some pre-requisites that we need to download or install i.e., anaconda environment, python and a code editor. **Anaconda**: Anaconda is like a package of libraries and offers a great deal of information which allows a data engineer to create multiple environments and install required libraries easy and neat.

To check on how to install anaconda environment, set up jupyter notebook refer to this article. You can skip the article if you have knowledge of installing anaconda, setting up environment and installing requirements.txt

1. Install necessary libraries from requirements.txt file provided.



1. Go to the directory where your requirement.txt file is present.

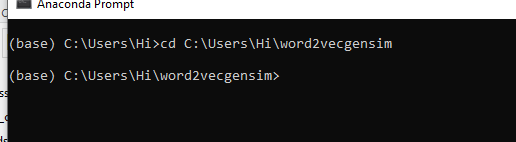
cd <<directory of your file>>. E.g, If my file is in d drive, then

1. cd d:

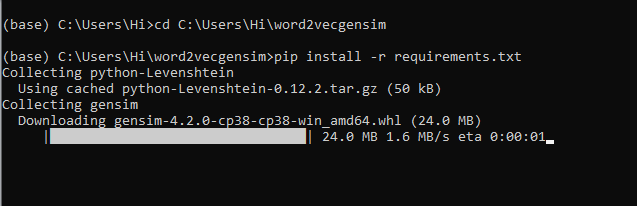
2. cd d:\License-Plate-Recognition-main  **#CHANGE PATH AS PER YOUR PROJECT, THIS IS JUST AN EXAMPLE**

If your project is in c drive, you can ignore step 1 and go with step 2.

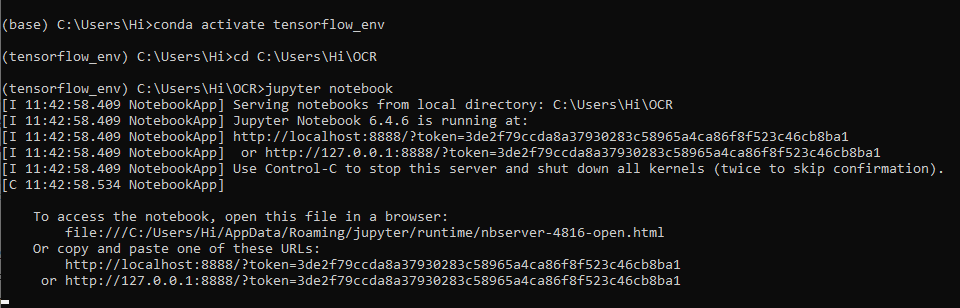
Eg. cd C:\Users\Hi\License-Plate-Recognition-main **#CHANGE PATH AS PER YOUR PROJECT, THIS IS JUST AN EXAMPLE**



1. Run command **pip install -r requirements.txt** or **conda install requirements.txt** (Requirements.txt is a text file consisting of all the necessary libraries required for executing this python file. If it gives any error while installing libraries, you might need to install them individually.)



All the necessary files will get downloaded. To run the code, open anaconda prompt. Go to virtual environment if created or operate from the base itself and start jupyter notebook, open folder where your code is present.

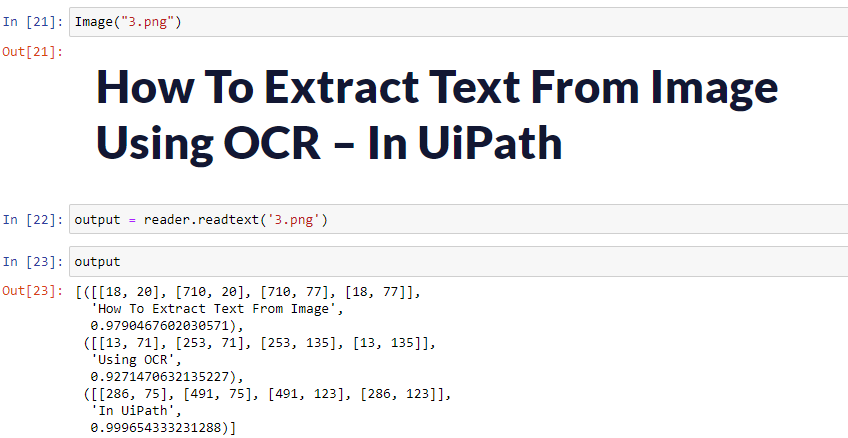




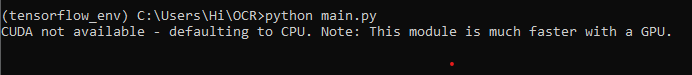
Open “easyocr.ipynb” to get the results.

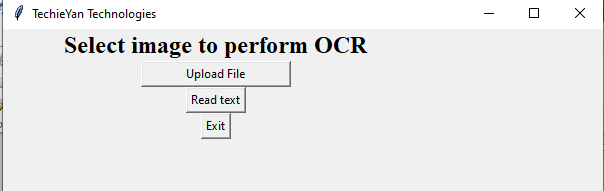
**Results:**

****

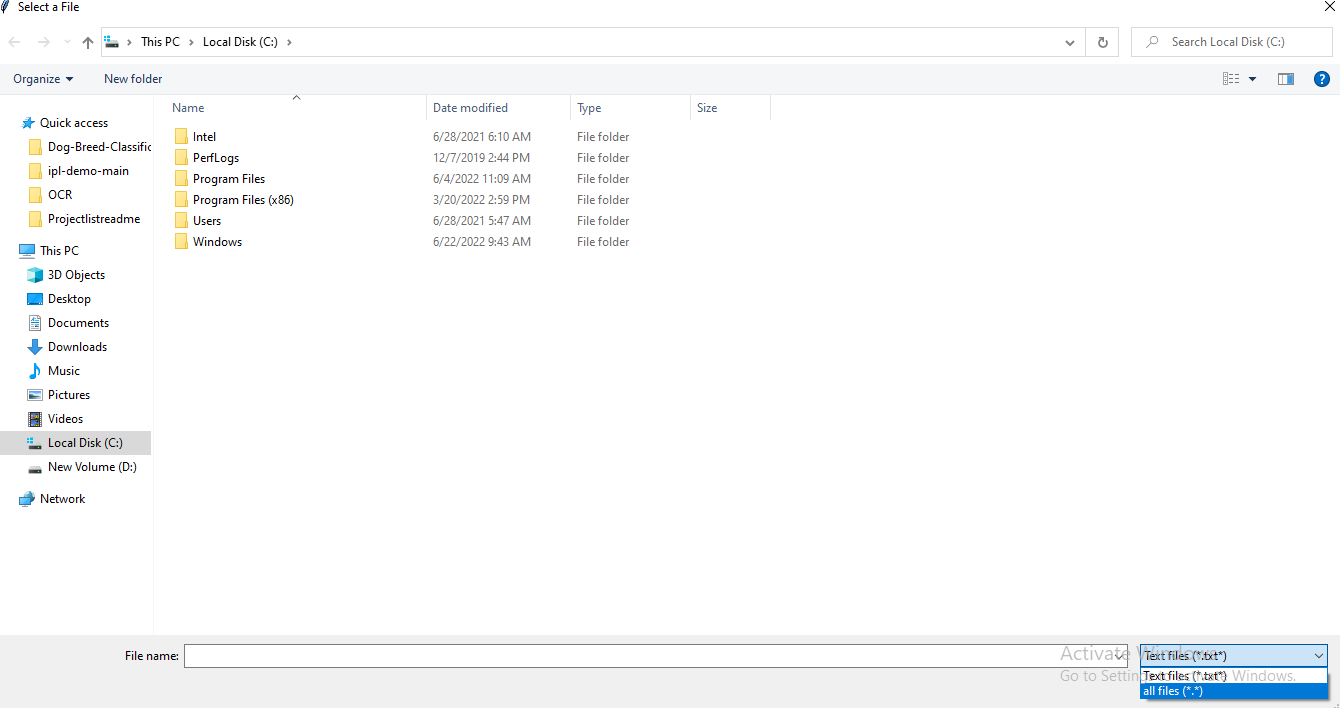
****

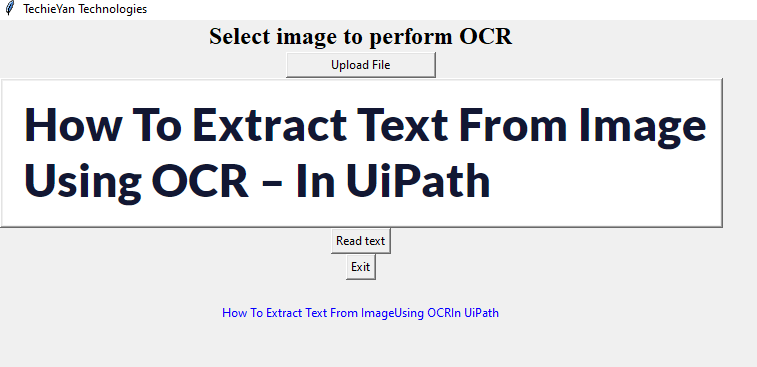
You can also run the “main.py” file to get the results. A tkinter window will be opened where you will have to select the file you want to perform OCR on and will get the results.





**Note:** While selecting files select option of all files, to be able to access images.



****

**Issues faced/ Points to note:**

1. Go to the current working directory (path of your project) to run main.py
2. Ensure you have all libraries installed.
3. EasyOCR can sometimes create version mismatches. If you get any errors, they will mostly be based on version mismatch. Browse the web and set it accordingly.