# Configuration Manual

MSc Research Project
Data Analytics


Aliasgar Wadhwanwala
Student ID: x21236381

School of Computing
National College of Ireland


Supervisor: Teerath Kumar Menghwar

**National College of Ireland**
**Project Submission Sheet**
**School of Computing**

| | |
|---|---|
| **Student Name:** | Aliasgar Wadhwanwala |
| **Student ID:** | x21236381 |
| **Programme:** | Data Analytics |
| **Year:** | 2023 |
| **Module:** | MSc Research Project |
| **Supervisor:** | Teerath Kumar Menghwar |
| **Submission Due Date:** | 14/12/2023 |
| **Project Title:** | Configuration Manual |
| **Word Count:** | 767 |
| **Page Count:** | 5 |

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

**ALL** internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

| | |
|---|---|
| **Signature:** | |
| **Date:** | 14th December 2023 |

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:**

| | |
|---|---|
| Attach a completed copy of this sheet to each project (including multiple copies). | ☐ |
| **Attach a Moodle submission receipt of the online project submission**, to each project (including multiple copies). | ☐ |
| **You must ensure that you retain a HARD COPY of the project**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. | ☐ |

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

| **Office Use Only** | |
|---|---|
| Signature: | |
| Date: | |
| Penalty Applied (if applicable): | |

# Configuration Manual

Aliasgar Wadhwanwala
x21236381

## 1 Introduction

This manual is intended to provide a comprehensive guide to installing and configuring the software application known as Python, Google Colaboratory, Roboflow and Jupyter. This guide will walk you through the entire process from start to finish, covering topics such as system requirements, installation, configuration, and troubleshooting. Additionally, it will provide helpful tips to ensure that your installation and configuration are successful.

## 2 Hardware Details

Hardware is the physical components of a computer system, such as the central processing unit (CPU), memory, storage devices, network cards, motherboards, and other components. Below are the details used in this project in figure 1. Local environment has been used in the initial phase of the project to sort the data and its visualization if any.

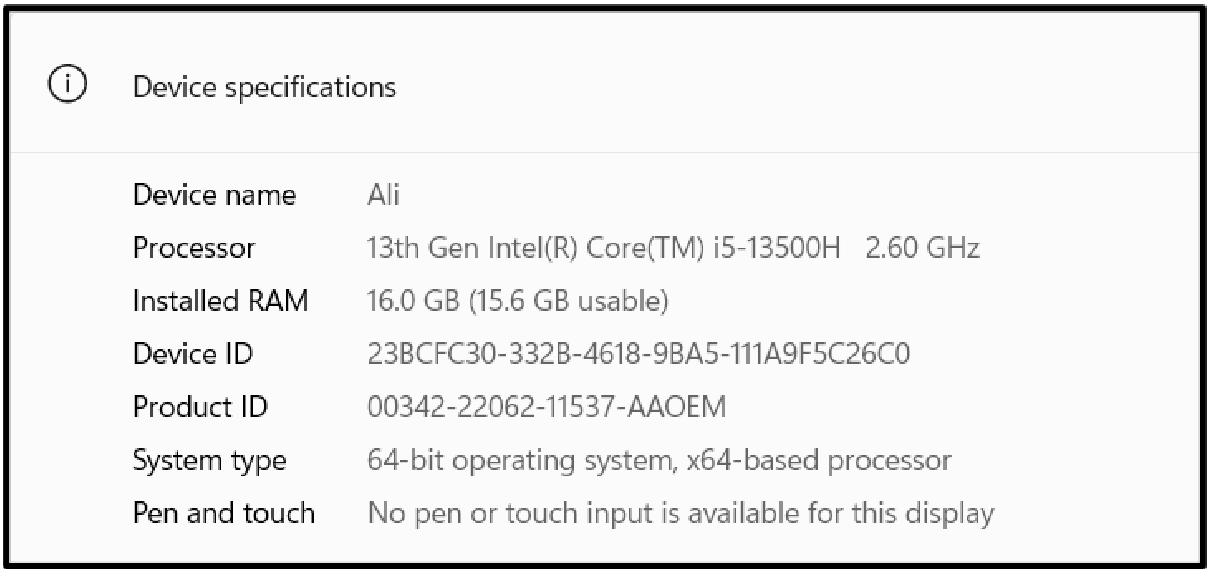


Figure 1: Hardware Details

# 3 Software Details

We have made use of Google Colaboratory and Roboflow software in this research. Other tools used will be described below.

- Programming Language - Python 3.10.12
- Cloud IDE - Google Colaboratory
- Annotation Tool - Roboflow[1]

## 3.1 Google Colaboratory

1. Sign-in to your Google Account and go to Google Colaboratory.
2. Create a new notebook or open an existing notebook that has been created earlier.
3. In the Google Colaboratory Toolbar, as shown in figure 2, select Runtime -¿ Change Runtime type and choose the type of instance you require.

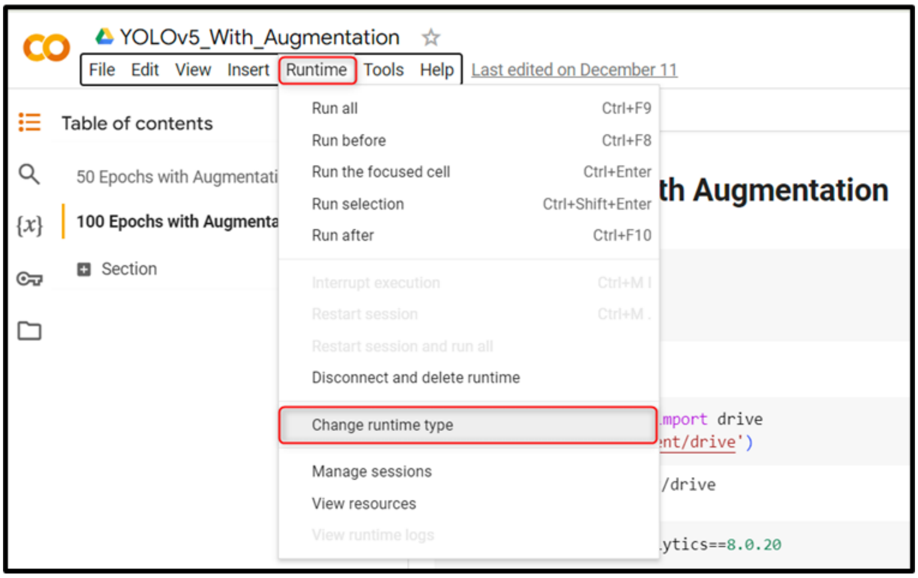


Figure 2: Changing Runtime

4. Select T4 GPU and click save, as shown in figure 3.
5. We need to import few packages that will help in the changing directories as and when required, as shown in figure 4.
6. Next we need to connect our Google Drive to the Google Colaboratory application, which can be done using the code in figure 5.
7. We have to install Ultralytics Package to train our YOLOv8 Model in Google Colaboratory environment and some checks need to be done to ensure all dependent packages have been installed, as shown in figure 6
8. This will complete the initial setup of the environment required for running our notebook.

[1] https://app.roboflow.com/

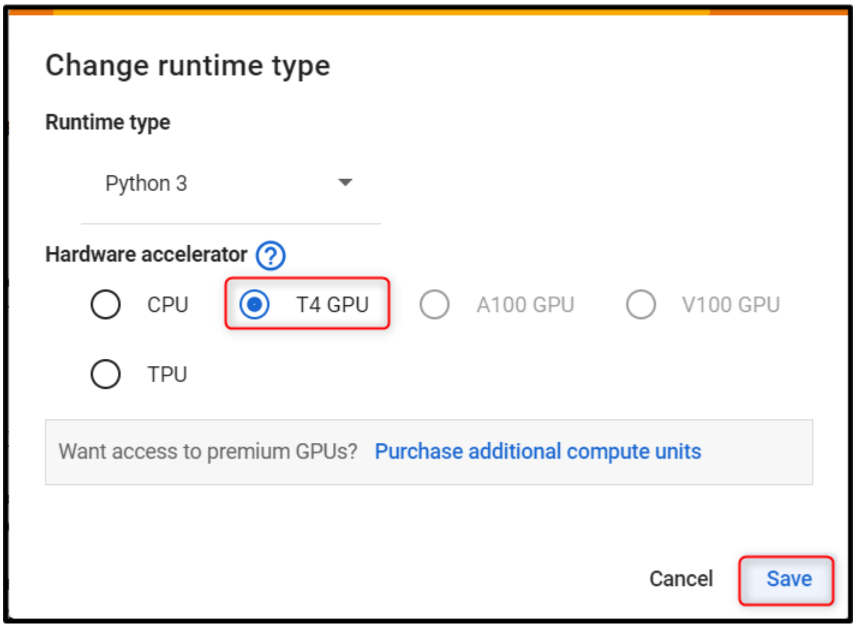


Figure 3: Selecting T4 GPU

```
import os
HOME = os.getcwd()
print(HOME)

/content
```

Figure 4: Directory Package

```
[ ] from google.colab import drive
    drive.mount('/content/drive')
```

Figure 5: Google Drive Connection

```
[ ] !pip install ultralytics==8.0.20

    from IPython import display
    display.clear_output()

    import ultralytics
    ultralytics.checks()

    Ultralytics YOLOv8.0.20 🚀 Python-3.10.12 torch-2.1.0+cu118 CUDA:0 (Tesla T4, 15102MiB)
    Setup complete ✅ (2 CPUs, 12.7 GB RAM, 26.9/78.2 GB disk)
```

Figure 6: Ultralytics Package installation

# 4 Dataset

The Dataset has been manually scraped from the web for images of waste objects and has also been combined with the Taco[2] dataset and has then been uploaded to Kaggle[3].

[2]http://tacodataset.org/
[3]https://www.kaggle.com/datasets/aliasgarwadhwanwala/waste-objects-in-wild/data

The dataset consists of a total of 970 images which have then been pre-processed before augmenting them.

## 4.1 Dataset Pre-Processing

The images collected have been then annotated using a tool known as Roboflow[4]. The image to be annotated has to be uploaded on their webapp and can be seen in the figure 7. We then have to create bounding box as required covering the whole of the object being annotated, in this case a plastic bottle. Next, the appropriate class has to be selected and you can move on to the next image for annotation as seen in figure 8

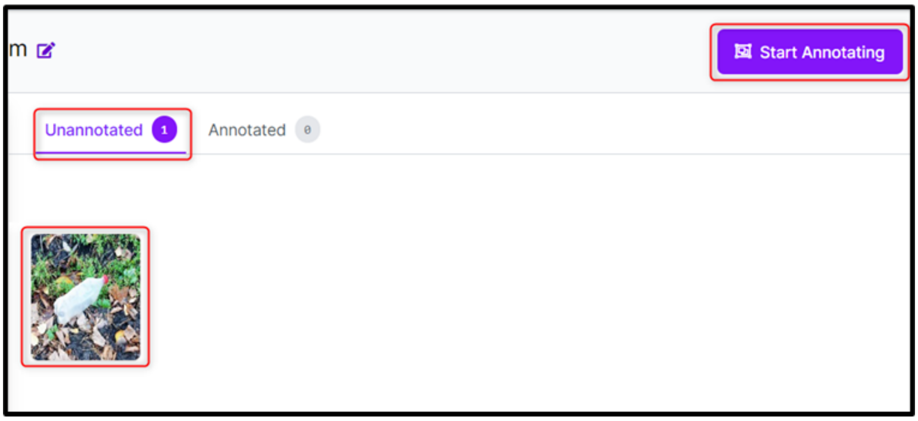


Figure 7: Image to be annotated

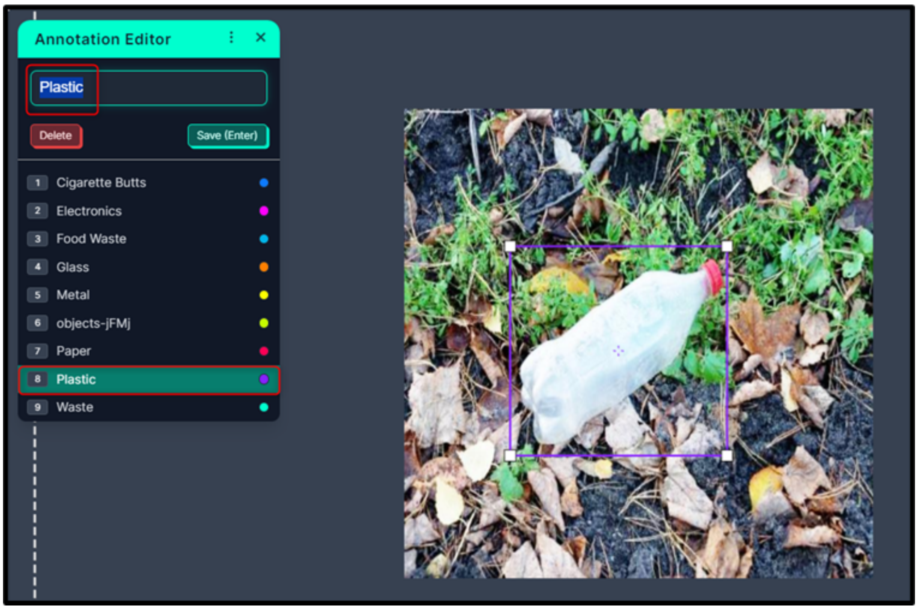


Figure 8: Annotating an image with 'Plastic' class

After all images are annotated we will then split the images in random into training, validation and testing sets. Their ratios are 80:15:5 respectively.

[4]https://app.roboflow.com

## 4.2 Data Augmentation

We have augmented our data in our notebook which can be seen in figure 9. As observed, we have used Adam optimizer, batch size of 32, the images have been rotated by 90 degrees and flipped vertically and horizontally. The number of epochs have been changed for different iterations for comparisons.

```
#Training

%cd /content/drive/MyDrive/Waste_Dataset

!yolo task=detect mode=train model=yolov5mu.pt data=data.yaml plots=True imgsz=640
lr0=0.001 optimizer='Adam' batch=32 epochs=50 degrees=90 scale=0.2 shear=0.2
flipud=0.25 fliplr=0.25 workers=8
```

Figure 9: Augmentation used during training

# 5 Modelling

This research required multiple iterations of training our model using different and epochs and with or without augmentation. In figure 10, we can see that a 50 epoch run without augmentation for a Yolov8 model required around 1 hour. Iteration with 100 epochs required almost double the time required than the 50 epoch run.

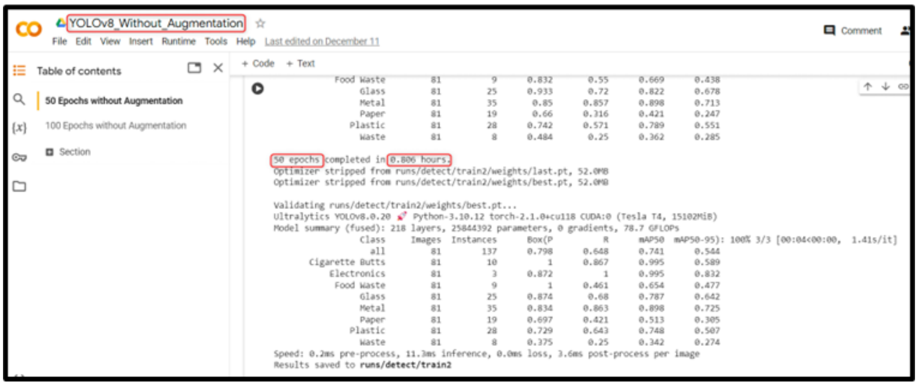


Figure 10: Time required for a single run