

Configuration Manual

MSc Research Project MSc in Data Analytics

Tarunveer Subhash Shetty Student ID: x19238754

> School of Computing National College of Ireland

Supervisor: Dr. Mohammed Hasanuzzaman

National College of Ireland Project Submission Sheet School of Computing



Student Name:	Tarunveer Subhash Shetty
Student ID:	x19238754
Programme:	MSc in Data Analytics
Year:	2021
Module:	MSc Research Project
Supervisor:	Dr. Mohammed Hasanuzzaman
Submission Due Date:	31/01/2022
Project Title:	Configuration Manual
Word Count:	989
Page Count:	8

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.

<u>ALL</u> 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:	Tarunveer Subhash Shetty
Date:	30th January 2022

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):	

Identification and Classification of Industrial Plastic Waste Using Deep Learning Models

Tarunveer Subhash Shetty x19238754

1 Introduction

This document will list the hardware and software platforms on which the research was conducted. The best effort has been made to list down the setup and execution steps undertaken in a logical sequence that can be easily repeatable. In this research, three models were built for the classification of plastic waste using python. These models were then compared to find the best working model and compared with the state-of-the-art models.

2 System Cofiguration

This section will list down the hardware configuration of the system and the software setup that helped in meeting the research goals.

2.1 Hardware Cofiguration

A personal computing device was used for the research work and the configuration of the system is shown below in Table 1:

Hardware	Configuration
System	Dell Inspiron-7572
System Type	64 bit
RAM	8 GB
Graphics	GeForce MX150
SSD Memory	128 GB
HD Memory	1 TB
Processor	Intel i5-8250 U

 Table 1: Hardware Configuration

2.2 Software Configuration

The software setup including the operating system and environment setup have been listed in this section.

The operating system details are listed in Table 2:

Specification	Value
Edition	Windows 10 Home Single Language
Version	20H2
OS Build	19042.1348

 Table 2: Operating System

For this research due to limited system resource availability, Google Colab was picked as the choice of platform for building the python code. The Colab platform just requires a standard browser to open and access. The browser used for this research was Google Chrome as it provides support to the Colab platform with high integration capability due to the availability of extensions on the browser. The version of Chrome in use is shown in Figure 1.



Figure 1: Browser Specification

2.3 Google Colab Configuration

The following steps need to be followed for opening a Google Colab notebook as shown in Carneiro et al. (2018) to run python:

- First you would need a active google account
- Go to https://colab.research.google.com/?utm_source=scs-index
- Sign in using your google account credentials. You should be able to view the following page as shown in Figure 2
- From here you can start building your python code by opening a new notebook

CC	Welcome To Colaboratory File Edit View Insert Runtime	Tools	Help						cə Share 🛱	T
≔	Table of contents	×	Examples	Recent	Google Drive	GitH	lub (Jpload	Connect 👻 🖍 Editing	^
Q	Getting started									
	Data science		Filter notebooks		Ŧ					
~~	Machine learning					Local contract of the second	-	-		
{ <i>x</i> }	More Resources		Title			Last opened 🔺	First opened 👻			
-	Machine Learning Examples		CO Welcome To Cola	aboratory		4:17 PM	4:17 PM			
	Section									
									to Colab to learn more, or	
									ets you write and execute	
									a result:	
									, room,	
							New notebook	Cancel		
			To execute the code	in the above cell, select	: it with a click and then	either press the	play button to the l	eft of the cod	e, or use the keyboard	

Figure 2: Google Colab

2.4 Google Colab Cloud Configuration

The environment hosted by Google Colab has limitations on the resource utilization when executing the python codes. This was leading to delays in the code execution times. To overcome this a google cloud instance had to be setup with high resource configuration to power the Google Colab notebooks. The steps to be followed for the Google Cloud VM instance setup as shown in Bisong (2019) have been listed below:

- First you would need a active google account
- Go to https://cloud.google.com/
- Sign in using your google account credentials. You should be able to view the following page as shown in Figure 3
- Go to the console page
- In the search box, search for "Colab"
- Go to Colab and launch it as show in Figure 4
- Make a selection of the machine type and deploy a Colab instance. The details of the configuration used for the research are tabulated below in Table 3

Instance Setting	Configuration
Machine Type	n1-highmem-2
CPU	Intel Haswell
GPU	Nvidia Tesla T4
SSD Memory	200 GB

Table 3:	Instance	Configuration
----------	----------	---------------

- Deploy the instance and keep it running
- Go to Google Colab and connect the notebook to GCE VM instance and provide the details of the Colab VM instance created. This allows the Colab notebook to access the high computational power of the Google Cloud VM



Figure 3: Google Cloud Homepage

=	Google Cloud	Platform	🖫 My First Project 👻				
4							
		Colab					
	CO	Google Colab					
		A hosted no	otebook solution for machine learning				
		LAUNCH	VIEW PAST DEPLOYMENTS				
		DDICINC	SUDDODT				
	OVERVIEW	PRICING	SUPPORT				

Figure 4: Google Cloud-Colab Launch

3 Project Implementation

3.1 Data Collection

The WaDaBa dataset by J. Bobulski (2018), has been used in the research and data was downloaded from the official WaDaBa website ¹. The dataset is of 4 GB and is available to be downloaded. The download is available for 20 sets of images with images types evenly divided in each set.

	W A D A B A
<u>Home</u>	Plastic Waste DataBase of Images – WaDaBa
<u>Database</u> <u>description</u>	
<u>Samples</u>	The main purpose of this work was creation of a plastic waste database of images of objects constituting the typical
<u>Term of use</u>	using methods of Computer Vision. Digital images of items that will be processing should reflect the specific conditions of
<u>Download</u>	places where real objects will be found. Thus, each thing is placed in this database should be presented in the course of several collections of images, taking into account different lighting conditions and different arrangement relative to the
Pet recognition	image recorder, and the different degree of deformation. Images created in the collection divided into groups based on the type of material from which individual objects were made.
Contact	
	Recycling is one of the most important methods of environmental protection, the purpose of which is to reduce the amount of waste stored in landfills and conservation of natural resources. The term recycling means the recovery of raw materials, which consists in the transformation of substances or materials contained in waste in the production process to obtain the substance or material for reused or other purposes, including organic recycling, excluding energy recovery. In

Figure 5: WaDaBa Database

3.2 Data Preparation

The image files present in the WaDaBa database are not annotated, labelled or classified in separate folders. There was need to manually label the image files by understanding the naming convention followed for each plastic type. The manual annotation of the files was performed on Roboflow online tool 2 .

¹ WaDaBa plastic database: http://wadaba.pcz.pl/

² Roboflow: https://app.roboflow.com/login



Figure 6: Roboflow Annotation

. The tool helps with several annotation features, after annotations the files are available to download in the desired format as shown in Figure 7.

🕜 roboflow	Projects	Universe	Documentatio	n											
		Ima			~							+ 4	dd Images	Gene	er
() - <mark>0</mark>		Tra	aining Set 280	o Vali	dation Set	801	Testing Se	et (400)	Unannota	ated 0					
				N	1			Ń	• 1		5 ()		2	K	
TARUNVEERS.PACKT@GMAIL.0	COM E	- ¹		1000				*		ø			•		
Overview		4	2.						THE R						
Annotate	0				-		ELOSE								
Dataset	2				1	-	0								

Figure 7: Roboflow Dataset

3.3 Data Preprocessing

The data splitting and augmentation have been performed in the Roboflow app. The dataset has to be split into train and test set for efficient model training and testing. Roboflow also offers a set of augmentation functionalities on the images. The dataset is later downloaded in the MultiClass Classification format which offers the data categorized into different folders based on the type of plastic as shown in Figure 8.

roboflow Projects	Universe Documentation				
	Plastics RP Data	aset			
	Generate New V	Version 2021	-12-07 6:47pm		
	VERSIONS	Export		×	Exp
'ARUNVEERS.PACKT@GMAIL.COM	2021-12-07 6:47pm v2 Dec 7, 2021 Experiment_1 v1 Dec 3, 2021	Format Multiclass Classifi Converts your object of	cation letection dataset into a classification dataset.	~	
Overview					
上 Upload		Cancel	Contin	ue	
Annotate 0		IIVIAGE			
Dataset 2		9321 ii	mages		

Figure 8: Roboflow Dataset Export

3.4 Model Building

Python is our base for building the code and all models in the research will be built in python. Python needs to import certain packages and install them before the models can be initialized as the model definition needs to be imported from the different libraries. The major libraries that need to be installed and imported have been listed below:

- Numpy- Used to create objects like arrays, matrices that support mathematical computations for the models
- Pandas- Used to create objects like series, dataframes that support data cleansing and analysis for the models
- Keras- Used to import the objects like model, layers, metrics, applications, and so on
- TensorFlow- Used to import the objects like models, layers, metrics, optimizers, applications, callbacks and so on

3.5 Model Evaluation and Visualizations

The models built are trained and tested with the train and test data split. The comparison of the models is necessary to conclude with the best results. There libraries that offer great power in terms of visualizations and evaluations. The major libraries that need to be installed and imported have been listed below:

- sklearn- Used to import objects like confusion_matrix and classification_report
- matplotlib- Used to import plot packages that offers plotting features that help in visualizing the results

References

- Bisong, E. (2019). Building machine learning and deep learning models on Google cloud platform: A comprehensive guide for beginners, Apress.
- Carneiro, T., Da Nóbrega, R. V. M., Nepomuceno, T., Bian, G.-B., De Albuquerque, V. H. C. and Reboucas Filho, P. P. (2018). Performance analysis of google colaboratory as a tool for accelerating deep learning applications, *IEEE Access* 6: 61677–61685.
- J. Bobulski, J. P. (2018). Pet waste classification method and plastic waste database wadaba, Advances in Intelligent Systems and Computing 681: 57–64.