Configuration Manual

MSc Research Project
MSc Data Analytics

Palash Shinde
Student ID: x19218681

School of Computing
National College of Ireland

Supervisor: Noel Cosgrave

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Palash Shinde
Student ID:	X19218681
Programme:	MSc Data Analytics **Year:** 2020 - 2021
Module:	MSc Research Project
Lecturer:	Noel Cosgrave
Submission Due Date:	23/09/2021
Project Title:	Marine Debris Segmentation using Capsule Network (SegCaps)

Word Count: **Page Count:**

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:	Palash Shinde
Date:	23/09/2021

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

Palash Shinde
Student ID: x19218681

1   Introduction

Data sources, software, system specifications, hardware specifications, tools, libraries, and code can be found in the configuration manual for the "**Marine Debris Segmentation using Capsule Network (SegCaps)"** project. Additionally, it describes how the data was processed before the implementation of the proposed models.

2   System Specifications

Detailing the configurations of the project's systems are presented in this section.

2.1   System Hardware to processes the data

Operating System:	Ubuntu 18.04
Processor:	Intel(R) Core(TM) i5-9300H CPU @ 2.40GHz 2.40 GHz
Installed RAM:	8.00 GB (6.48 GB is usable)
System Type:	64-bit operating system x64based processor
Hard Disk:	1 TB

2.2   System Hardware for the training of SegCaps deep learning model

It was becoming computationally expensive to perform training of model on the hardware mentioned in 2.1. Hence AWS EC2 instance (p3.2xlarge) is created. Instance configuration is specified below.

Operating System:	Ubuntu 18.04
RAM:	61GB
vCPUs	8
GPU	GPUs - Tesla V100
GPU Memory	16 GB

3   Softwares Used

Python programming language is used for the data preprocessing, building model and evaluation of the model on test data. Pip packages of numpy, scipy, tqdm ,scikit-learn, scikit-image, pillow, SimpleITK, tensorflow-gpu, keras, h5py is used for the project. [1]

These packages are listed and can be quickly installed with the help of the requirement.txt file (from code artefacts) by executing the below command in any python environment. The command is stated as per below.

```
$ pip3 install -r requirement.txt
```

4   Project Implementation

4.1   Data Acquisition & Preparation

Marine Debris Data is first sourced from TrashCan[2] portal. Dataset.zip file is downloaded from the web portal. This zip contains RGB images of marine debris belonging to 16 classes in total. Post unzipping of the file. Mask images are generated from the bbox annotations present in the JSON file of the dataset. The images (with7212 images) and annotations(with7212 annotations .json files) folder obtained from Dataset.zip should be unzipped and place in the same path as the dataset_prep.py. Along with this, another dir of material_version must be created in the same path of dataset_prep.py file. The following command helps in creating the mask images from the .py file (from the code artifacts folder)

```
$ python3 dataset_prep.py material
```

Post execution of the dataset_prep.py. This will create files in files in ./material_version/train/ & ./material_version/mask/

4.1.1   Model Training & Data Augmentation

For SegCaps model training main.py script is required. This script mainly requires a root path for the training & mask images. '--train' flag for the script must be set to '1' for using the script in training mode. To include Augmentation while training --aug_data must be set to '1'. The following command helps in initiating SegCaps model training on Marine debris data

```
$ python3 main.py --data_root_dir './material_version/' \
  --train 1 --aug_data 1
```

4.2   Model Evaluation on Test Data

For SegCaps model testing main.py script is required. This script mainly requires a root path for the test images. '--test' flag for the script must be set to '1' for using the script in testing mode. The following command helps in initiating SegCaps model testing on Marine debris images

```
$ python3 main.py -- test_imgs_dir './test_imgs/' --test 1
```

---

[2] https://conservancy.umn.edu/handle/11299/214865

4.2.1 Metrics Calculation

For Metrics calculation on SegCaps model main.py script is required. This script mainly requires a root path for the test images. '--test' flag for the script must be set to '1' for using the script in testing mode. Along with this --test_metric flag must be specified. If it is set to '1' dice & IOU metrics are calculated. Along with this if a flag is set to 2 precision as well as metrics both are calculated as well. The following command helps in initiating SegCaps model testing on Marine debris images & metrics calculation.

```
$ python3 main.py --test_imgs_dir './test_imgs/' --test 1 \
 --test_metric 1
```