Chapter 1

Maintenance Manual

1.1 Installation

To install the software, extract the sources from the archive available at our GitHub repository. Next, the compulsory Udacity lakeside and jungle map data, the preprocessed comma.ai dataset and the Udacity autonomous driving car simulator program, which were adopted for the project, can be downloaded at archive.org. Besides, the original, compressed 45 GB comma.ai dataset file is also available online [**?**].

1.1.1 Directory Structures

```
data
├── jungle_map
│   ├── IMG
│   └── driving_log.csv
└── lakeside_map
    ├── IMG
    └── driving_log.csv
```
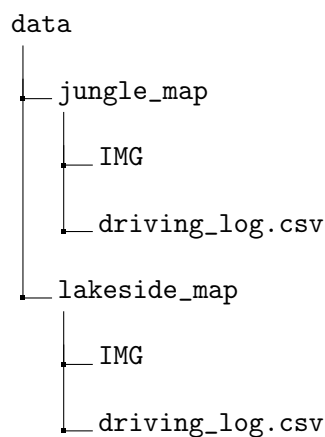
Figure 1.1: Udacity dataset directory tree. The 'IMG' subdirectories contain the recorded images. Conversely, the 'driving_log.csv' files store various measurement data such as steering angle, speed, throttle etc. related to each image.

```
dataset
├── 64_batched_cloudy
│   ├── angles
│   └── frames
├── 64_batched_night
│   └── ...
└── 64_batched_sunny
    └── ...
```
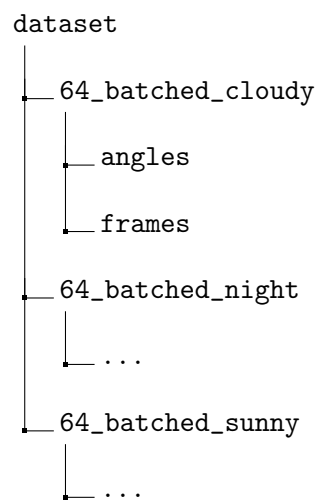
Figure 1.2: Comma.ai dataset directory tree. The 'angles' subdirectories contain the steering angle measurement data related to each image. On the other hand, the 'frames' folders comprise NumPy files each with 64 (batch size) driving frames.

1.2 Software Dependencies

All required program script packages and dependencies are set forth in Table 1.1 hereunder:

Package	Version
eventlet	0.30.2
Flask	1.1.2
h5py	2.10.0
Keras	2.4.3
keras-ncp	2.0.1
matplotlib	3.3.4
netifaces	0.10.6
numpy	1.19.2
pydot	1.4.2
python-dotenv	0.16.0
python-engineio	3.8.2
python-socketio	4.2.1
Pillow	8.1.0
Python	3.8.7
scikit-image	0.18.1
scikit-learn	0.24.1
seaborn	0.11.1
setuptools	54.2.0
sklearn	0.0
tensorflow	2.4.1
tqdm (optional)	4.58.0

Table 1.1: Used Python code file packages and their versions.

1.3 Temporary Files

The only temporary files created are the Keras ModelCheckpoint HDF5 ones that store the pre-trained version of the neural networks and are created at the end of each epoch. Throughout the project we emphasise solely the best MSE loss result version of the models. Hence, all the other epoch checkpoints can be omitted, particularly when training the models via unseeded variants of the scripts.

1.4 Requirements

1.4.1 Disk Space

The project program files are only couple MB in size, which every contemporary computer system hardware can handle comfortably. However, the bulk of the Udacity dataset is roughly 2 GB whereas downloading the full comma.ai dataset in raw format would require a minimum computer disk space of 45 GB.

1.4.2 Memory (RAM)

Processing of large datasets may involve holding a substantial amount of data in memory. Therefore, it is recommended that the program scripts are run on systems with at least 4 GB of RAM available.

1.5 System Files and Folders

In this section we explore the project directories by describing their content and summarising the role of the present source code files. Firstly, in the following directory tree, shown in Figure 1.3, we portray the contents shared by both model training directories. Next, we go over the dataset distinctive source files and folders in the subsequent sections.

```
udacity/comma.ai
│
├── models.py ...      Module that contains the developed model code defini-
│                         tions.
├── train_model.py ...    Script used to train the developed IL architectures.
│
├── evaluate_model.py ...    Model evaluation script.
│
├── utils.py ...      Module that defines general and widely used utility func-
│                         tions, which add functionality to the main source code files.
├── plot_datadis.py ...    Program file employed in plotting the steering angle data
│                             distribution of the comma.ai recordings.
├── images ...    Directory comprising various data processing and model
│   │                training and evaluation pictures.
│   ├── architectures ...    Subfolder with the terminal model summary and architec-
│   │                           ture plots.
│   ├── dist_plots ...    Steering angle data distribution plots directory of the lake-
│   │   │                   side and jungle tracks, and the augmented version of the
│   │   │                   former ones.
│   │   ├── terminal
│   │   └── plots
│   ├── prediction_plots ...    Predicted vs actual steering angle plots directory of the es-
│   │                             tablished neural nets.
│   └── loss_plots ...    Model training loss plots directory of the developed mod-
│                           els.
├── checkpoints ...    Folder that comprise pretrained HDF5 model files.
│   ├── cnn
│   │   ├── seeded
│   │   └── unseeded
│   └── ...
```
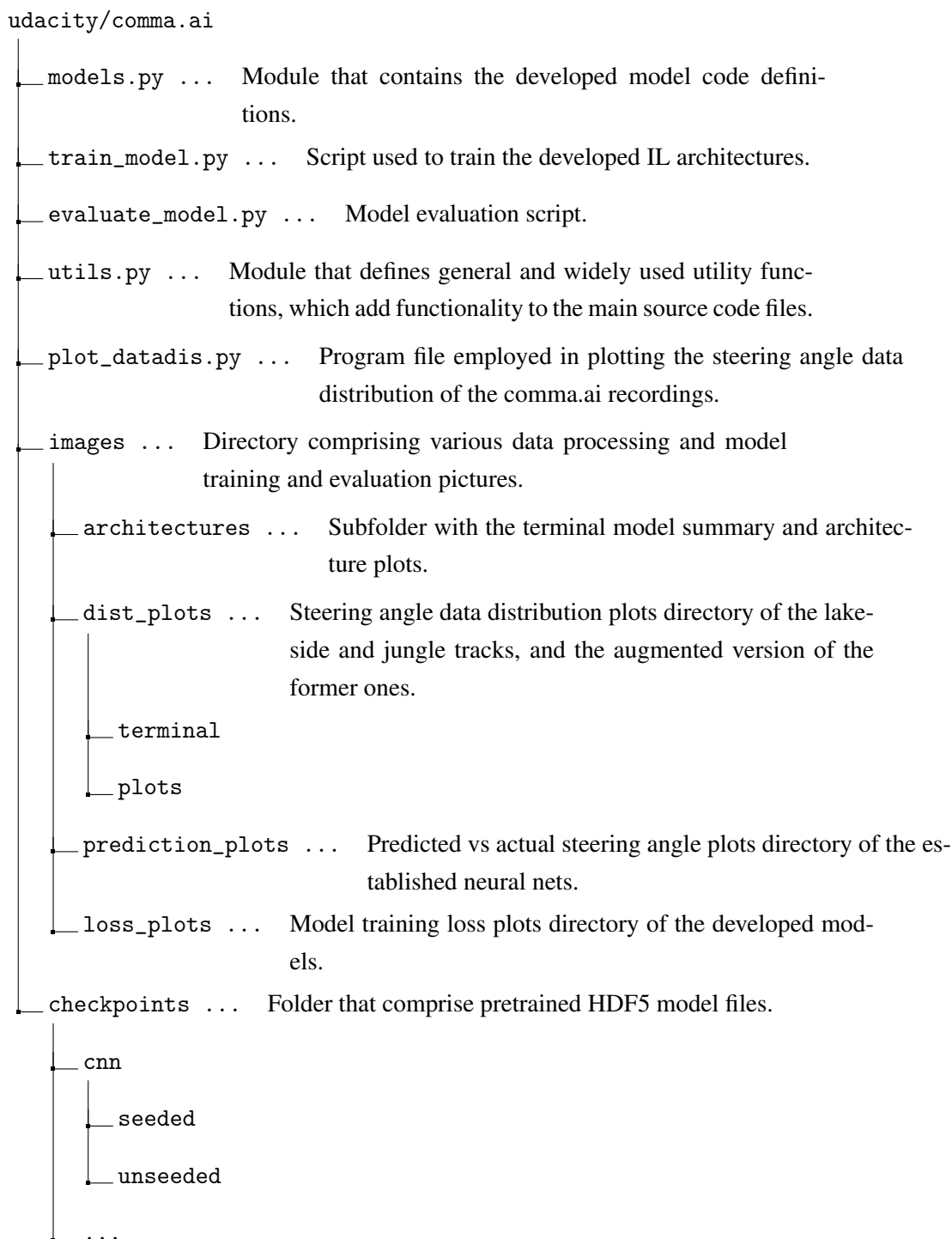
Figure 1.3: Shared files and folders.

1.5.1 Udacity Dataset

The Udacity dataset model training folder is named 'udacity'. In addition to the previously depicted files and folders, we specify the individual components of this directory in Figure 1.5 below.
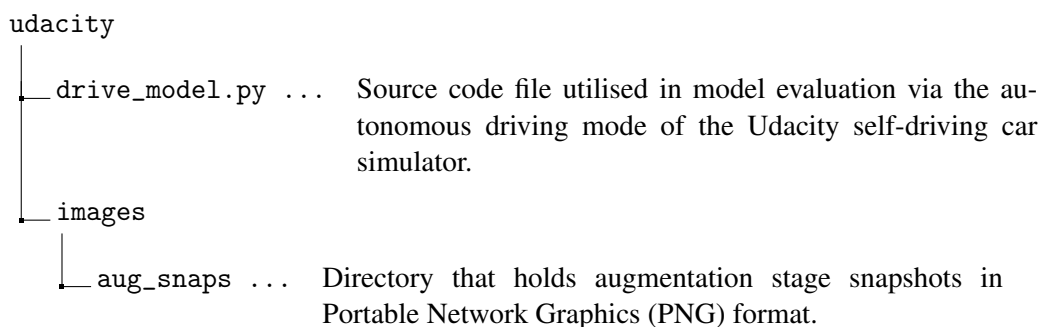
```
udacity
│
├── drive_model.py ...   Source code file utilised in model evaluation via the au-
│                        tonomous driving mode of the Udacity self-driving car
│                        simulator.
│
└── images
    │
    └── aug_snaps ...    Directory that holds augmentation stage snapshots in
                         Portable Network Graphics (PNG) format.
```

Figure 1.4: The specific Udacity model training directory files and folders.

1.5.2 Comma.ai Dataset

The folowing directory tree shows the comma.ai dataset specific files and folders on top of the aforementioned shared ones.
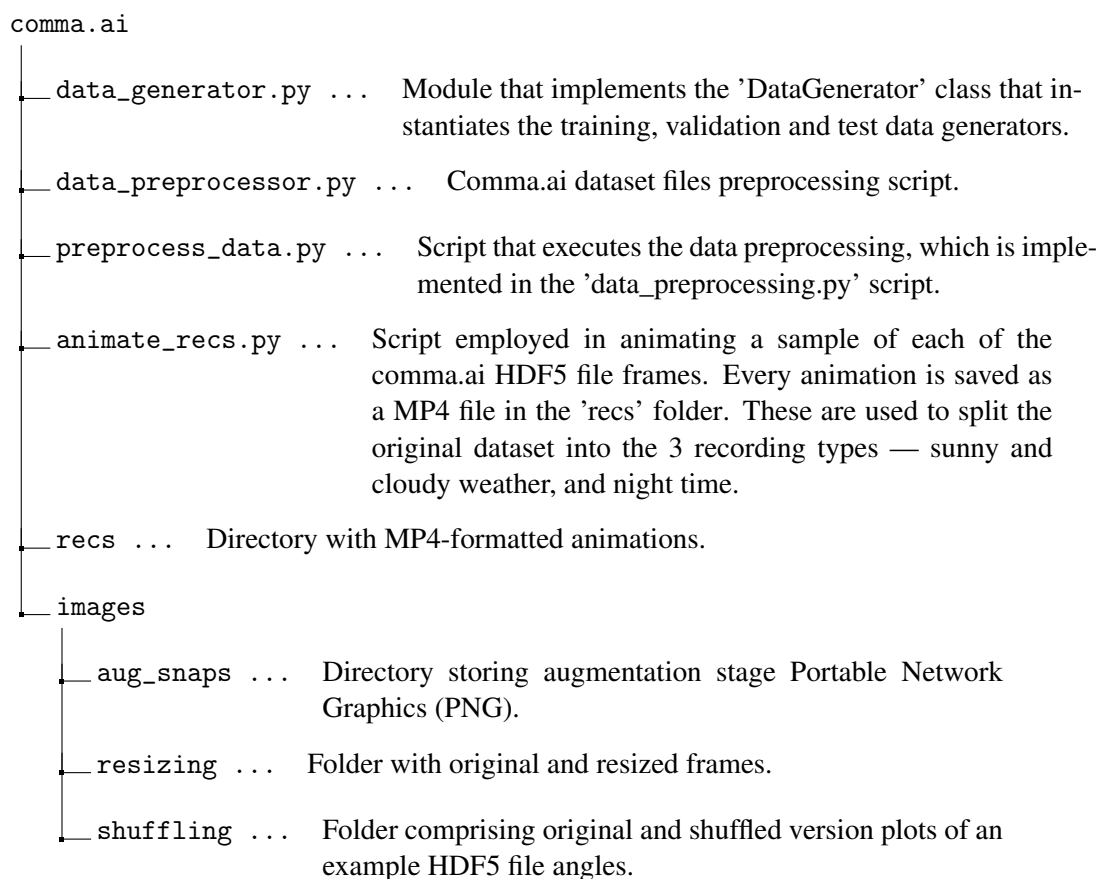
```
comma.ai
│
├── data_generator.py ...    Module that implements the 'DataGenerator' class that in-
│                            stantiates the training, validation and test data generators.
│
├── data_preprocessor.py ...   Comma.ai dataset files preprocessing script.
│
├── preprocess_data.py ...   Script that executes the data preprocessing, which is imple-
│                            mented in the 'data_preprocessing.py' script.
│
├── animate_recs.py ...   Script employed in animating a sample of each of the
│                         comma.ai HDF5 file frames. Every animation is saved as
│                         a MP4 file in the 'recs' folder. These are used to split the
│                         original dataset into the 3 recording types — sunny and
│                         cloudy weather, and night time.
│
├── recs ...   Directory with MP4-formatted animations.
│
└── images
    │
    ├── aug_snaps ...   Directory storing augmentation stage Portable Network
    │                   Graphics (PNG).
    │
    ├── resizing ...   Folder with original and resized frames.
    │
    └── shuffling ...   Folder comprising original and shuffled version plots of an
                        example HDF5 file angles.
```

Figure 1.5: The specific Udacity model training directory files and folders.