IT Carlow –

BSc.

Software Development

UAV using Convolutional Neural Networks

Functional Specification

Student Name: Josh Hudziak

Student Email: C00231846@itcarlow.ie

Date: 13/11/2020

Supervisor: Dr. Oisin Cawley

Supervisor Email: oisin.cawley@itcarlow.ie

Table of Contents

Table of Figures

Introduction

This document contains the system requirements, both functional and non-functional, for a UAV (unmanned aerial vehicle) using CNN (convolutional neural networks) to achieve autonomous navigation. This UAV will be able to navigate over a path, road, or track without the aid of a human. Within this document, topics such as motivation for creation, target market, metrics for usefulness and potential risks will be stated.

As the dataset we used will be categorical, as seen below, UAV navigation can be deemed as a classification problem. Therefore, the CNN that is used shall return a class prediction to the UAV that the probability the image it has received is a turn either left, right or the image is central along the path. Then, the UAV can move in a safe direction with this probability as it's context to the path direction.

The UAV cannot train itself with what technology it has, therefore a CNN is required. The CNN will train frame by frame, breaking each frame into layers of convolution and pooling. These layers will finally connect into a fully connected layer with a probability of three classification neurons; left, right and straight. From this information the UAV should be able to navigate a road, path, or trail.

In order to be able to train a model, a dataset is required which is labelled and big enough to support the model that is being designed. This dataset will be compiled of footage of driving on roads and lanes, as well as hiking on trails and paths. This generalization of terrain will make the UAVs navigation more functional in different situations. As this footage will be gathered in Ireland, an array of visual differences should appear in the footage, such as wet and dry surfaces, dark and bright lighting and visual imperfections from weather in general. This footage must be split into frames and labelled into three classifications; left, right and straight.

Next is the problem of communicating this information, live, to the UAV. As the Parrot Anafi is being used, this system will communicate from a laptop using Olympe which is a python-based controller interface for this brand of UAV.

Target Market

Due to the nature of autonomy which will be computer vision using CNNs to accomplish navigation of roads, paths and tracks, this UAV can be used in many industries for many purposes. Although the initial use case of navigation seems extremely basic, the technology used in the navigation task can be adapted for more complex issues. This is because the CNNs used in computer vision are dissecting image frames from the UAV and there is an enormous amount of useful information within images. All that must be done is for the algorithm to be adapted and trained for each use case.

This UAV technology is fantastic for search and rescue in concealed areas with paths where GPS cannot function. Due to its visual processing nature the UAV can search path areas. Of course, the UAV must be trained to recognise persons in distress. Although this only covers a specific area of search and rescue it still saves time for humans to look elsewhere. A report by Aerosociety has shown that civilian drones used in different ways have been saving 1 life per week on average [REA17].

A similar area of interest is first aid. A UAV would carry life saving medicine to a person in need. Navigating through streets and identifying persons in distress, the UAV can then release its payload. An article in IEEE Spectrum studying the response time of an ambulance and a drone found the drone to be faster at transporting medicine [SCU19]. The life saving medicines delivered by the drone could be the difference between life and death while an ambulance arrives.

This UAV can also be used in the agricultural sector. On some farms drones are used for different agricultural processes but this is in tandem with agricultural machinery and systems. Specifically, this UAV design with path navigation through computer vision can aid fields that tractors have been used on as they leave behind tracks. Therefore, the UAV can monitor the field with ease, this then opens more areas of specific agricultural use such as crop planting and spraying, soil and field analysis, irrigation and health assessment [MAZ16]. As with previous examples of targeted markets, the use case is not specific and flexible with modifications for each speciality of agricultural needs.

There are many other possibilities of use for a UAV using this specific system, such as street surveillance, delivery services and military uses. The power of vision-based navigation alone is an extremely powerful tool.

Metrics

Measuring success in the development stage is based upon; the quality of the acquired dataset, different deep neural network and CNN techniques that will be used in the system. This will be clearly seen in the testing of the neural network on the test dataset which it will never have seen. The performance tools used will be the root mean squared algorithm and the explained variance ratio. These tools will assess the model's accuracy of its current model after being trained on the training dataset against the unseen test dataset.

To achieve an autonomous UAV in these trial stages, a laptop will run Olympe and Sphinx software (created by parrot, the drone manufacturer) which in turn will communicate and direct the UAV. Thus, this connection must be stable and image processing must be received at a high framerate in real time to effectively run the CNN model. Therefore, the higher the frame rate received the greater the chances the model has at success.

Next a baseline must be achieved to recognise the model is functioning correctly. This baseline must be simple as to recognise the communications are working correctly as well as the model working correctly, thus, the first baseline metric will be for the UAV to navigate a straight road, track are path, and so, the first major physical milestone will be achieved.

Following the first baseline being achieved, the next major milestone will be for the UAV to navigate turning on a selected road, track, or path.

Figure 1: An example of the task to be completed for the first physical baseline achievement

Figure 2: A basic left-hand turn to achieve the second major physical achievement.

This should be a complete navigation from a straight path into a turn either left or right and continuing to go straight after the turn has been completed.

To fulfil the UAVs purpose, it must finally complete a circuit of straights and turns with different gradients and visual path textures, thus, achieving generalization of road, tracks, or paths. The navigation should happen in a seamless manner without flight stuttering and pausing as this would be inefficient.

Figure 3: A circuit with an outlined track. There are many different visual textures to process and many different turning angles to process for a final achievement.

Architecture

Due to the technical nature of this project there will not be a graphical interface for users but the person monitoring the drone will receive video feedback from the drone which will indicate what operation the drone is currently taking, this will be labelled left, right or centre. The project consists of a Parrot Anafi drone with a built in camera, the drone will communicate via Parrot Olympe framework receiving input from a Convolutional Neural Network hosted on a Dell XPS 15 using an Intel i7-9750 running at 2.6ghz – 4.5ghz. The XPS also has an Nvidia GTX 1650 which will speed up the process of training and testing the model rather than an integrated Intel GPU.

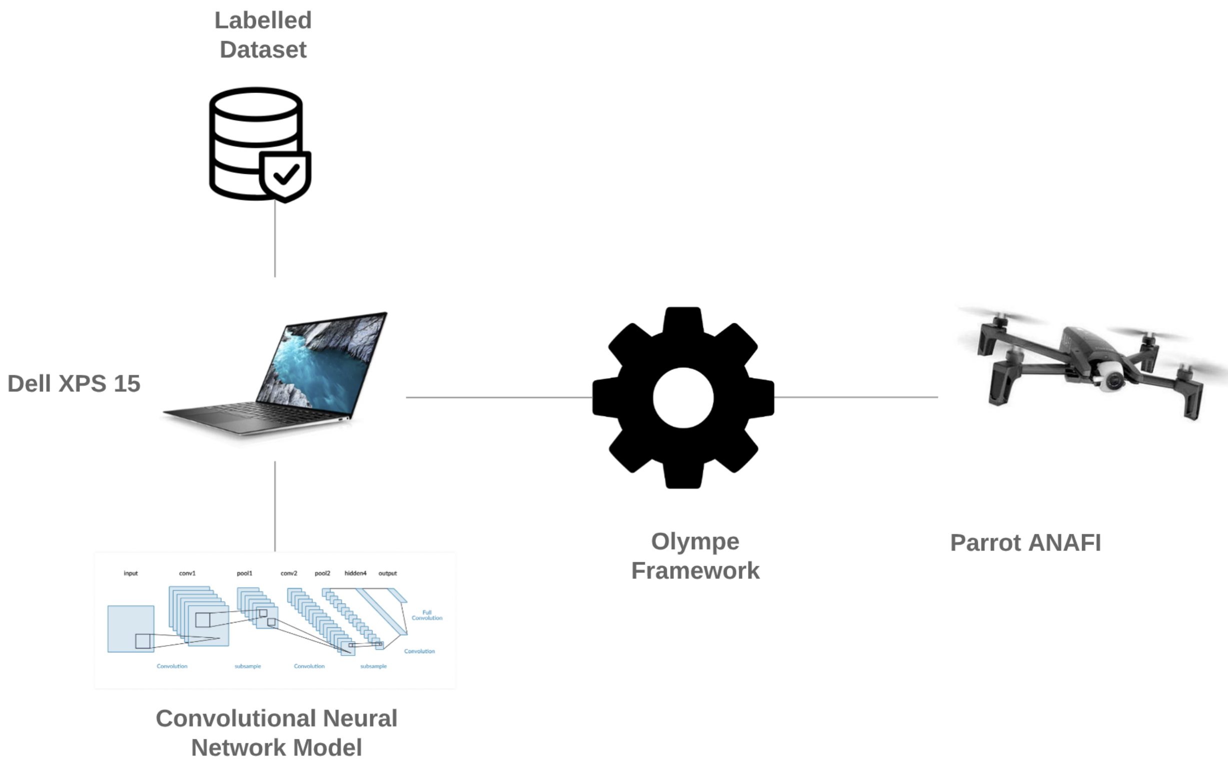

Figure 4: A basic overview of the project's architecture.

The CNN will be made using ResNet-18 architecture [HE16]. While this architecture follows the same convolutional setup as VGG [SIM15] or AlexNet [KRI12], it introduces the feature called skip connections. This skipping feature enables a model to become extremely deep with many layers. An identity from a previous layer can be passed over another layer acting like a link to past memory for the network. This identity converges with the input for the next layer and continues in this path. The skips can be used over blocks of layers either, these blocks may contain convolution, normalization, activation and dropout functions.

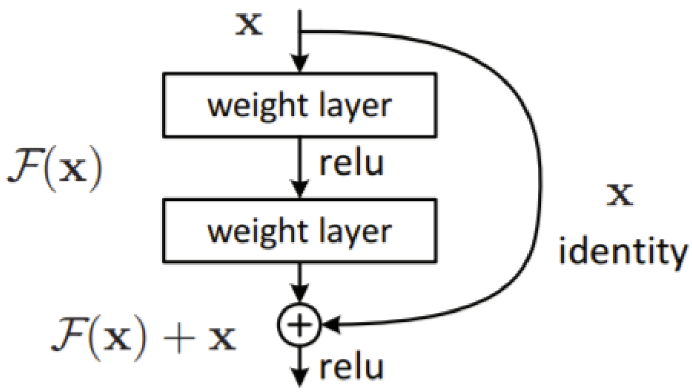

Figure 5: A residual learning block. Source: [HE16]

For the conventional convolutional layers a small filter size will be used such as (3x3). Each residual block in the network will be subject to batch normalization which will standardise the inputs and this will be followed by the Relu activation function to eliminate negative values within the input for each convolutional layer. Each Convolutional layer will have a property that can change the learning rate which is called a regulariser.

For each block the dimensions of the model will increase, to ensure regularization during this, the skip connection will pass through a (1x1) layer of convolution which does not add parameters, therefore, computational complexity remains the same.

The amount of residual blocks will vary with the amount of data available and the quality of this data. This is to ensure overfitting does not occur (model remembers training data but performs badly on tests).

The model will be connected using a flattening function which turns it into a one dimensional array. A dropout function can be applied here to overcome overfitting. Next the results are passed into the dense network using the softmax function.

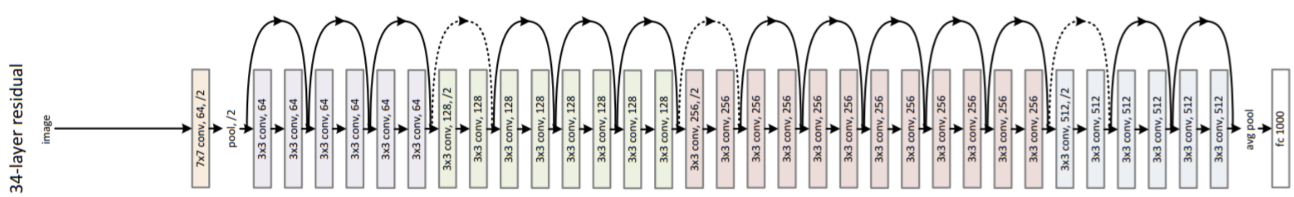

Figure 6: ResNet with 34 layers. Source: [HE16]

Technology and Tools

Many different libraries and technologies were researched but ultimately the following technologies and tools are being used to implement the project. The UAV that is being used is the Parrot Anafi and because of this the Olympe software must be used to control the drone. This software can only be run on Ubuntu 18.04. Olympe is also written in Python, as a result Python is the language of choice for this project.

Olympe

A programming interface that communicates with the Parrot Anafi. This program will act as a controller to the UAV and for simulated UAVs in Sphinx. Olympe connects to the UAV using Python scripts, thus, the user can send commands and receive statuses, information, and errors from the UAV.

Sphinx

Another Parrot software that simulates environments to test created features for the UAV. Sphinx runs Parrot's firmware and uses a software called Gazebo to produce virtual environments and a UAV simulation to test functionality and programs on. Sphinx can produce flight data and virtual behaviours.

Python 3.4

The base language for all development on this project. Python is used in this project as many of the other tools pair well and use python. Python has built in garbage collection and is very easy to implement. It is widely used in machine learning and has many libraries and frameworks for this such as scikit-learn.

TensorFlow with Keras

These libraries have many machine learning features built in and will simplify implementing the ResNet/ReLu and entire CNN. TensorFlow is an open source library built by Google's brain team. It holds a large repository of neural network algorithms and models which entail mathematical solutions such as convolution. Keras is a high-level API used for TensorFlow. It creates an easy implementation method for neural networks by using a block method.

Open CV

OpenCV is used to process images from the collected dataset. Using NumPy, this library has a repository of over 2500 algorithms to detect and track objects, thus solving a lot of problems within the computer vision field. We will be using it with Python, but it is also compatible with many other languages such as C++ and Java.

NumPy

NumPy is a library that works on large multi-dimensional arrays, it is extensively a Python library. It holds a vast repository of high-level mathematical algorithms. Using NumPy can generate results 50 times faster than normal Python lists. In this project we will use it in tandem with OpenCV to process images in the dataset.

Supplementary Specifications

Here the software quality attributes are reviewed. From this review a validation of the most necessary components can be made.

Functionality

These are the main features of the project.

- The CNN model outputs an image classification that is either left, right or straight.
- The UAV can move left, right or straight.
- The UAV can navigate a complete circuit unaided by a human.

Usability

As this project mostly focuses on research of convolutional neural networks in UAVs, there is not much user experience other than running the model and turning on the drone.

- The drone should be fully activated within 45 seconds from pressing the on button.
- A User of this project should be able to run the model within 1 minute using the command line.

Reliability

Reliability adheres to this project's accuracy and recoverability.

- Recovery from failure should take between 1.5 and 2 minutes when the UAV is within physical range of the user.
- The UAV should navigate a road, path, or track without issue 90% of the time.
- The model should run without issue 99% of the time.

Performance

These are the standard for the project's throughput of information and response times.

- Transmission of signals to the UAV should be in real time 99% of the time.
- The UAV should move at a consistent pace.
- The UAV should move steadily without jerking or wavering.
- The UAV activates autonomous navigation mode within 10 seconds of the CLI prompt.

Supportability

These are the factors which will make the project maintainable and scalable.

- The code base should be organized and commented for maintainability and to allow for future changes to be made easily.
- This project is not supportable without Ubuntu, specifically version 18.04.
- Users must have installed TensorFlow, Keras, NumPy, OpenCV, scikit-learn and Python 3.4.

Timeline Plan

Phase One

The first phase of this project will take place from November 27th to January 8th.

Work Scheduled

- Take footage of different roads, tracks and paths using the camera of a One Plus 7 pro and an Akaso Brave 7 action camera.
- Label frames from all the footage with 3 categories; left, right and straight.
- Develop the architecture needed and run it within the Sphinx simulator.
- Get the Olympe software fully functional in both Sphinx and on the Parrot Anafi.

Phase Two

The second phase of work will be carried out from January 15th to February 26th.

Work Scheduled

- An implementation of the dataset should be used on the model developed so far. This would include training on the dataset.
- Process the results from training and evaluate the dataset.
- Assess the accuracy of the data set using the root mean squared algorithm and the explained variance ratio.
- Further develop the model and make changes where necessary.

Phase Three

The third and final phase of development will be carried out from March 1st to April 2nd.

Work Scheduled

- Prepare the final model.
- Train and then test the model on the dataset acquired for the final time.
- Test the UAVs navigational skill on different types of roads, tracks, and paths.
- Film the results of different tests to prepare for presentation.
- Prepare Final report for submission.
- Create a screencast for submission.

References

[HE16] He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[KRI12] Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*, pp.1097-1105..

[LOD17] Lodaya, T. 2017. *keras-signs-resnet*. Retrieved from Github: https://github.com/tejaslodaya/keras-signs-resnet/blob/master/README.md

[MAZ16] Mazur, M., 2016. *technology review*. Retrieved from MIT Technology review:

https://www.technologyreview.com/2016/07/20/158748/six-ways-drones-are-revolutionizing-agriculture/

[REA17] Read, B., 2017. *News*. Retrieved from Aerosociety:

https://www.aerosociety.com/news/life-saving-drones/

[SCU19] Scudellari, M., 2019. *The Human OS*. Retrieved from IEEE spectrum:

https://spectrum.ieee.org/the-human-os/biomedical/devices/drone-vs-ambulance-drone-wins

[SIM15] Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Plagiarism Declaration

Declaration

- I declare that all material in this submission e.g. thesis/essay/project/assignment is entirely my/our own work except where duly acknowledged.
- I have cited the sources of all quotations, paraphrases, summaries of information, tables, diagrams or other material; including software and other electronic media in which intellectual property rights may reside.
- I have provided a complete bibliography of all works and sources used in the preparation of this submission.
- I understand that failure to comply with the Institute's regulations governing plagiarism constitutes a serious offense.

Student Name: Josh Hudziak

Student Number: C00231846

Signature: Josh Hudziak

Date: 30-04-21