ReLOaD – Reinforcement Learning for Active Object Detection

Alex Nicholson, Jen Jen Chung and Liang Wang

Thesis Aim

- To create an agent that can intelligently manoeuvre a mobile robot equipped with a camera and object detection to maximise the amount and overall accuracy of object detection (OD) data collected.

Background

- Visual inspection of a set of target objects is a very commonly required robotics task.
- Inspection robots must gather the best quality observation data given variable environmental constraints such as limited robot battery or time-available on site.
- ReLOaD pilots an object detection robot in the SimpleSim environment to maximise information gathering, adapting to these variable constraints.

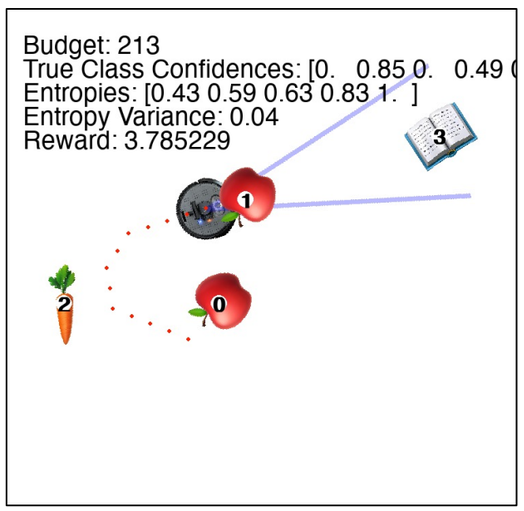

Fig. 1 – SimpleSim Environment

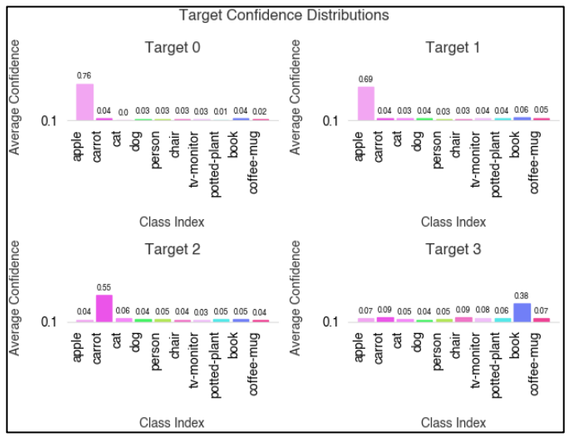

Fig. 2 – Confidence Histograms

Approach

- ReLOaD is an RL trained agent that controls the 2D movements of a non-holonomic robot.
- It aims to maximise the total information gain (creating a single strong peak in the confidence distributions of each class) given variable constraints – by planning the most 'informative' path through the environment.
- Trained in the custom SimpleSim environment (continuous state and action space) with simulated object detection confidences.
- Extensive experimentation was undertaken to tune the contents of the state and action space to achieve the best learning performance.

Outcomes

- Developed SAC-based RL model that adapts, without retraining, to changing time-budget conditions and number of targets
- Model successfully maximises the amount of information gain close to optimal performance by efficiently using the time available
- Developed custom SimpleSim simulator for training agents with simulated object detection inputs
- Agent trains in ~4hrs on Nvidia GTX4070 (4M iterations)
- Achieved 1.4k reward with an average maximum available of 2k (only achievable id the agent was able to instantaneously gain all available information)

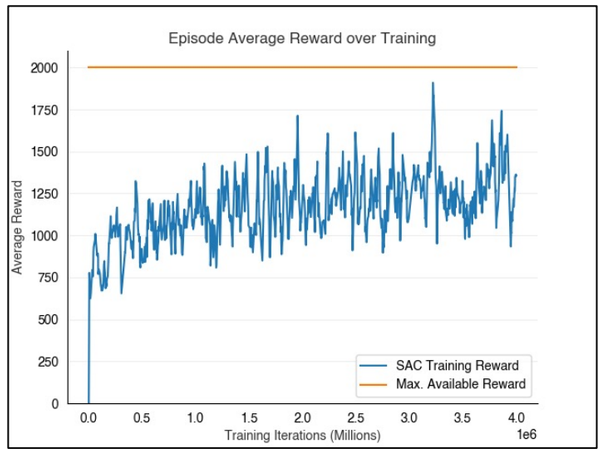

Fig. 3 – Training Episode Mean Reward

System Architecture

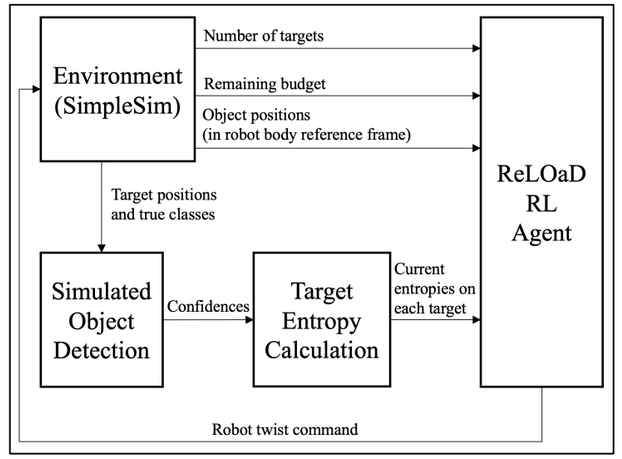

Fig. 4 – Reward Function

- Simulated object detection confidence is generated as a function of distance (closer is better) and viewing angle (viewing from front is better)

$$reward = R_s = \frac{1}{M}\sum_{m=0}^{M-1} IG_m$$

$$IG_m = \frac{1}{T}\sum_{t=0}^{T-1}\left(1 - \sum_{n=0}^{N-1} -p_n \log_N p_n\right)$$

where M = number of targets,
T = number of timesteps so far,
N = number of classes

Fig. 5 – Reward Function

School of Electrical Engineering & Computer Engineering

RPPL Lab Thesis

The University Of Queensland
AUSTRALIA