Machine learning based object tracking behavior for underwater vehicles

Abir Khan, Supervisor: Ingrid Schjølberg

abirk@stud.ntnu.no, ingrid.schjolberg@ntnu.no

Introduction

Recent researches has introduced Machine Learning (ML) into the robotics community in order to improve the robots' autonomous capabilities based on acquired experience. Reinforcement Learning (RL), which is one of ML methods, has especially attracted increasing interest rapidly. Basically, the RL system learns by interacting with the environment, then changes its policy regarding action selection by evaluating its feedback signals, thus, converging towards its desired behaviour.

In this thesis, a contribution to increased autonomy is made by creating a tracking behavior for ROVs. Essentially, the tracking behavior is of interest in this thesis as this implies that the need of an extra operator is not needed. Consequently, this technology saves the industry lots of expenses as it is not dependable on an extra operator, which is expensive and unavailable at times. Even though, in this thesis the ROV is learned to track a QR-code, the tracking-object of interest can easily be replaced by any other object of interest. For example, another ROV, thus, creating a master-slave system.

Objective and scope

The scope of work is the following:

1. Literature review on ML with emphasis on RL.

2. Design of a RL based tracking control for the ROV. The ROV is prompted to track a QR-code underwater.

3. Training of the RL model and performance check of the ROV with the use of simulator.

4. Implementation of the trained RL based control system on a real-world lab experiment.

Method

The following has been done so far:

1. An extensive **literature review** has been done. Where the fundamentals of ML (specifically RL) are introduced. State of the art RL algorithms are also presented in light of continuous environments. This is highly relevant for vehicular locomotion such as ROVs. The literature review is concluded with prevailing challenges in RL industry.

2. An online state of the art **implementation** of **Proximal Policy Optimization (PPO) algorithm** is found and checked for verification before being modified for use in this thesis. The line of thought is documented with the basis on knowledge from the literature review.

3. A **ROV simulator** is designed, with the physical hydrodynamic properties. This is where the ROV's RL based tracking behaviour is trained before being exported for use in physical lab experiment.

4. Perform **training** of the ROV itself with the implemented PPO algorithm. Present the line of thought in every critical choices made. Present the simulation results and discuss the ROV performance.

5. A **computer vision** algorithm is used to detect the QR-code from the ROV's camera. This is based on the widely used **OpenCV** library.

The following are being worked on:

1. Conduct physical **lab testing** and extract results. Discuss the performance. Detail the further work proposals. Conclude the thesis.

Implementation of PPO algorithm and simulator environment

The RL algorithm used in this thesis is build upon a powerful computational tool within data science called **neural networks**. Figure 1 (left) depicts how a neural network is structured. This is the backbone of the learning mechanism in the implemented RL algorithm (i.e. PPO). Figure 1 (right) shows how the simulator and RL algorithm interacts. Which is, in essence, a feedback based setup where the simulator feeds the RL algorithm with data (i.e. ROV states, terminal commands, reward etc) and the RL algorithm utilizes the data to train the model (i.e. ROV).
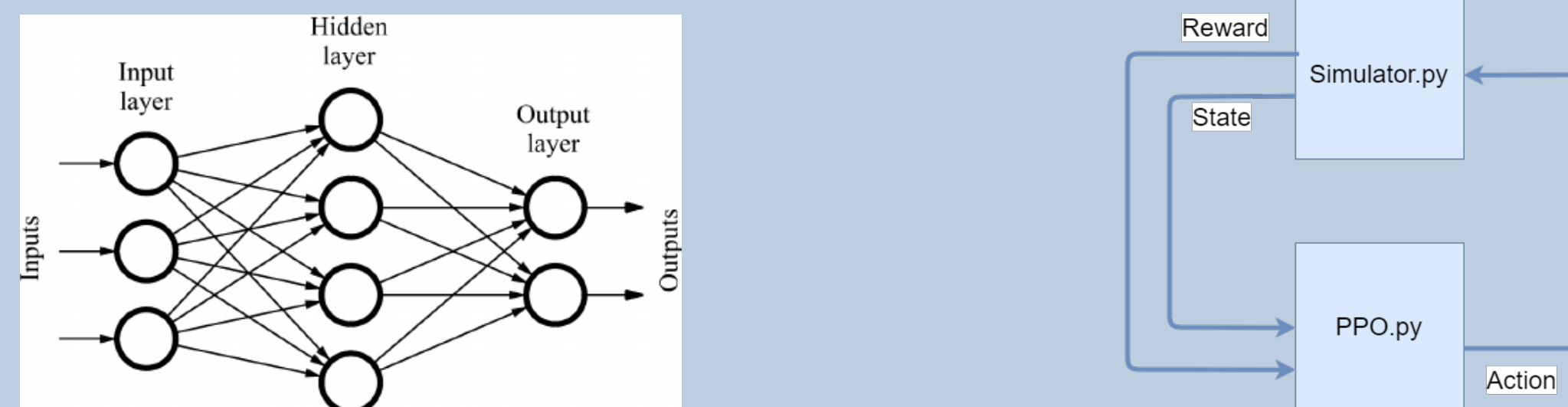

Figure 1: Neural network structure (left). RL-simulator setup (right)

The states which are interesting in light of RL algorithms are the body-fixed velocity of the ROV, the relative states. In this thesis, only planar environment is considered, thus, are the ROV's relative state to the QR-code the radial position r and relative yaw ψ. Figure 2 (left) shows how the relative states are determined during simulations. While figure 2 (right) shows the environment the ROV is enclosed in during training. The test basin used for physical experiments is the MC-lab at Tyholt, NTNU, thus, is the dimensions in the simulator basin correspondingly. The blue area in the right hand side figure is the ROV's deployment area during each training episode. This is to train the ROV to encroach the origin (where the QR-code position) is based on the relative states and its body-fixed velocity. Initial relative yaw ψ is randomly initiated during each training episodes as to simulate the many random relative states which real-life cases will yield during the experiments.
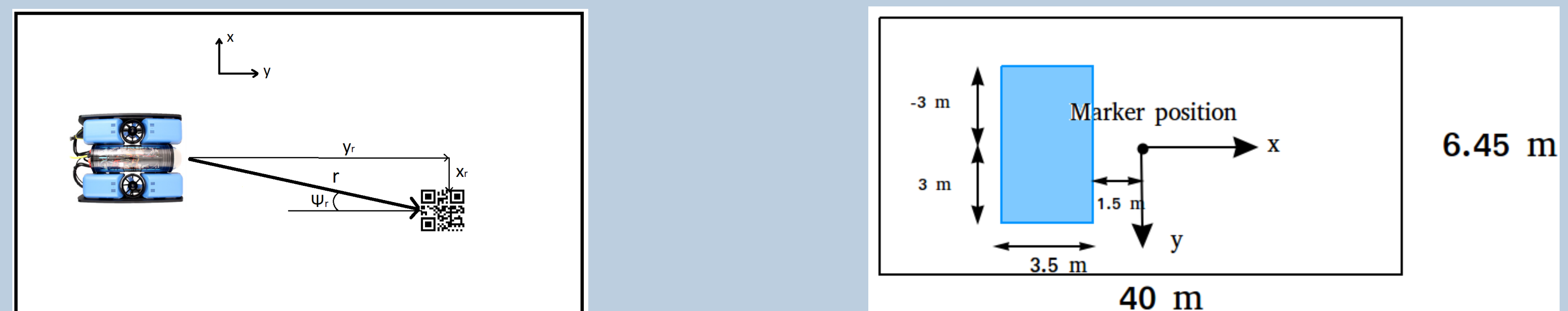

Figure 2: ROV - QR-code relative states (left). Simulator environment (right)

Simulation configuration and results

Before conducting physical experiments it is a common practise to verify the robot's performance in a simulator beforehand. By obtaining feasible performance in simulator indicate that there is a higher chance of achieving positive results in real-life experiments. Figure 3 the two arbitrary trajectory plots of the ROV. Each from different simulation episode. The red dot in the figure is the starting position, while the blue line is the trajectory history, and the green dot is the QR-code's position. The ROV is given 30 seconds to train before the episode is terminated, this is why the figure on the left has not reached the QR-code yet.
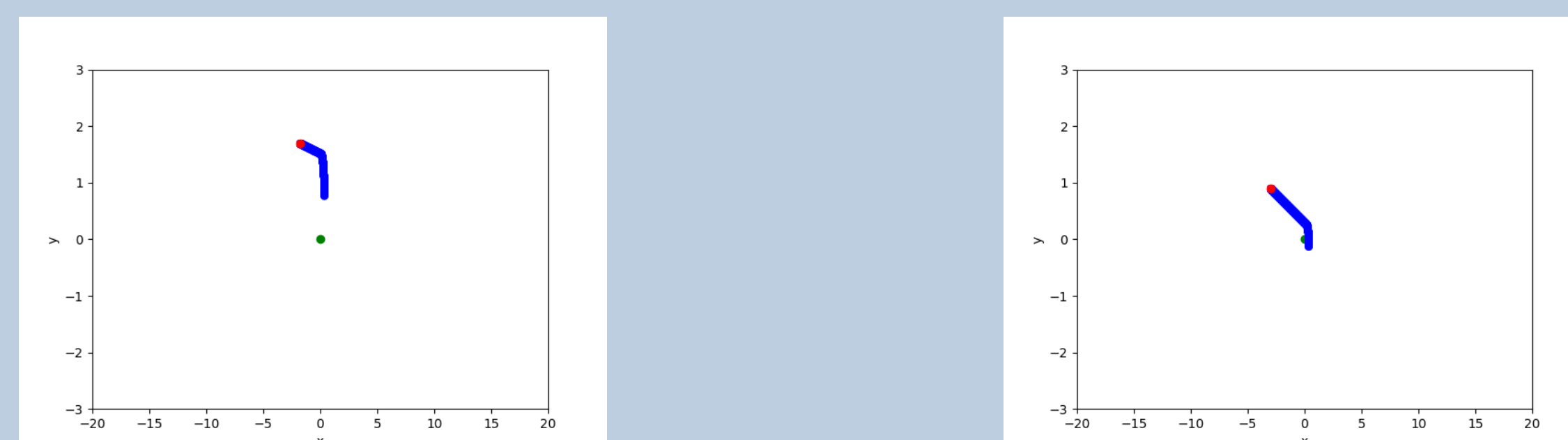

Figure 3: Trajectory plots from simulator-trained model.

The resulting lab testing stage is currently an on-going process. Figure 4 shows what the ROV "sees" from its camera view, using the powerful computer vision tool.
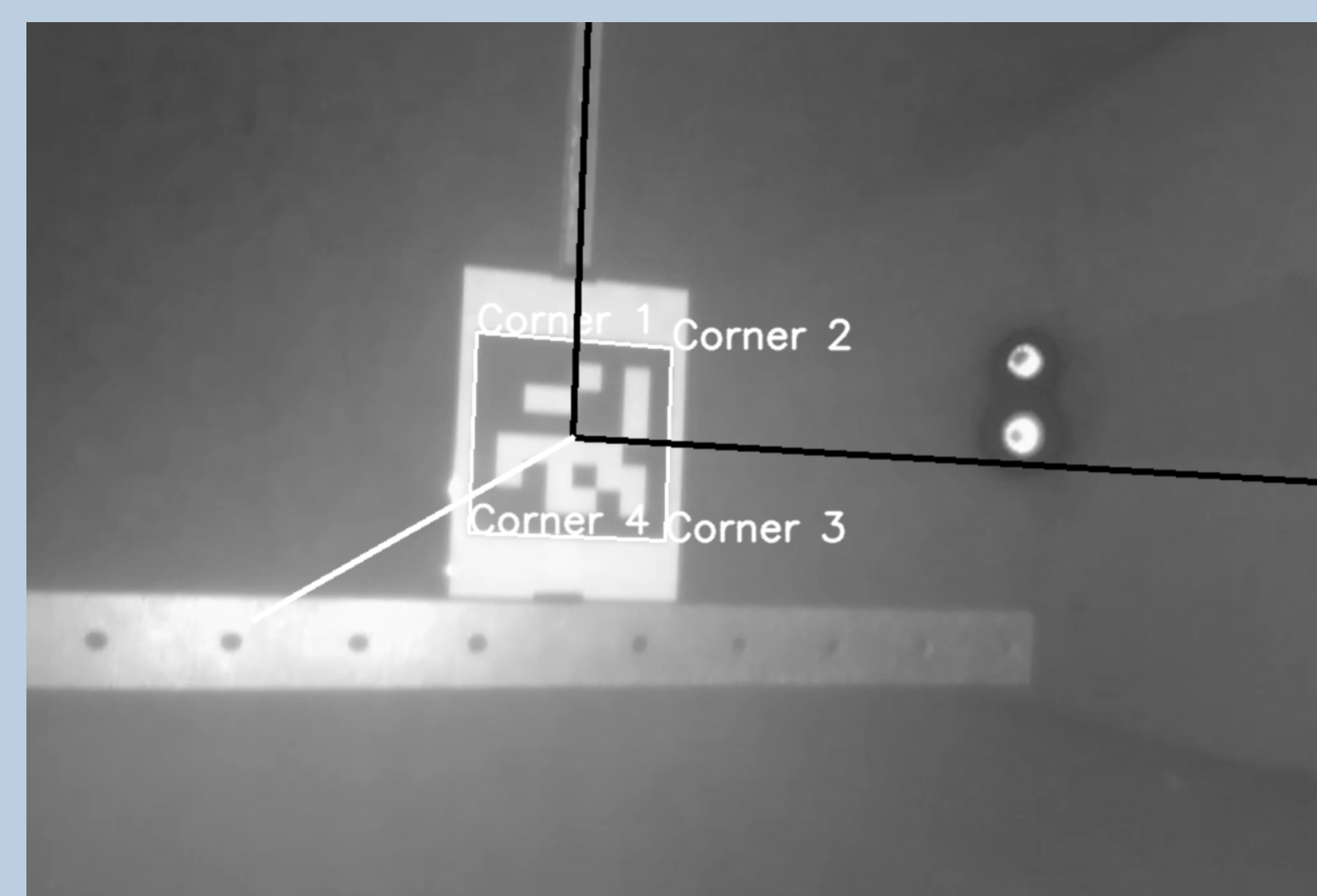

Figure 4: Trajectory plots from simulator-trained model.

Conclusion

The simulation results yielded promising performance, which remains to be imported to the real ROV and tested. While there are a lot challenges yet to remedy, the resulting real-life performance poses contains a modest chance of successful behavior. However, this thesis is a good starting point for further works and the beginning of a new way of increasing autonomy in the marine technology industry.