Fraunhofer
FKIE

# Tracking and Evasion using Co-Training with Context Knowledge

**Keywords**

Machine Learning, Reinforcement Learning, Co-Training, Path-Planning, Tracking, Evasion

## Description

Reinforcement Learning (RL) is a popular tool of Machine Learning that trains policies to solve potentially complex tasks and is already widely employed in the field of robotics. Applying RL to competitive environments, such as the boardgame Go or the computer game StarCraft 2 can happen by either training the policies using expert knowledge, or by allowing the policy to play against itself and gradually improve this way. While expert knowledge is typically prohibitively expensive, self-play on the other hand only requires longer training time. In addition, the trained policies are not biased to counter expert players, which results in more general policies and has already led to the discovery of new strategies in the boardgame Go.

In previous work, we have applied this concept to the field of target tracking and evasion. The simulation contains a UAV and a ground-based platform, both of which are controlled by an autonomous policy. While an observing platform has the goal to track a target with its radar, the target platform is supposed to evade this tracking as well as possible. This leads to conflicting objectives for the two agents, which are trained to dynamically counter each other. This work has addressed simple, empty environments and an extension to a more complex environment may lead to even more interesting behaviours of the agents.

## Goal

The goal of this thesis is to extend the existing Co-Training approach to an urban environment. This environment should include buildings and roads, which will be passed to the (RL-)agents as scenario-specific context knowledge, e.g. using maps. Consequently, the observation model of the agents needs to be extended with a simple CNN to utilize this context information. Furthermore, different Reinforcement Learning architectures, such as model-based RL or Attention Networks may be explored. In the end, the Co-Training should result in an equilibrium between the two conflicting agents, where both agents utilize the unique features of the urban environment to achieve their respective task. The final evaluation should include a comparison to the model without context knowledge, such that the advantages of the approach are clearly outlined.

## Requirements

- Solid Python programming skills
- Basic knowledge about Reinforcement Learning
- Some experience with NumPy and TensorFlow
- Basic Linux command-line knowledge may help

## Reference Paper

A. Brandenburger, F. Hoffmann and A. Charlish, "Co-Training an Observer and an Evading Target", *2021 IEEE 24th International Conference on Information Fusion (FUSION).*