FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

STUDENTSKÁ VĚDECKÁ KONFERENCE 2022

Analysis of the state of table football and prediction of its change based on image data

Bc. Matěj Sieber[1]

1 Introduction

The task is to plan positions for defenders and goalie to create a simulator for shooting practice in real arena. The chosen approach to this task is reinforcement learning, due to its potential for surpassing supervised learning methods. To successfully learn the agent a simulation environment in ROS in combination with Gazebo and 3D model in Autodesk Fusion 360 was made. In the real arena camera was used to estimate player position and rotation as well as ball position. The 3D model can be seen in Figure 1.

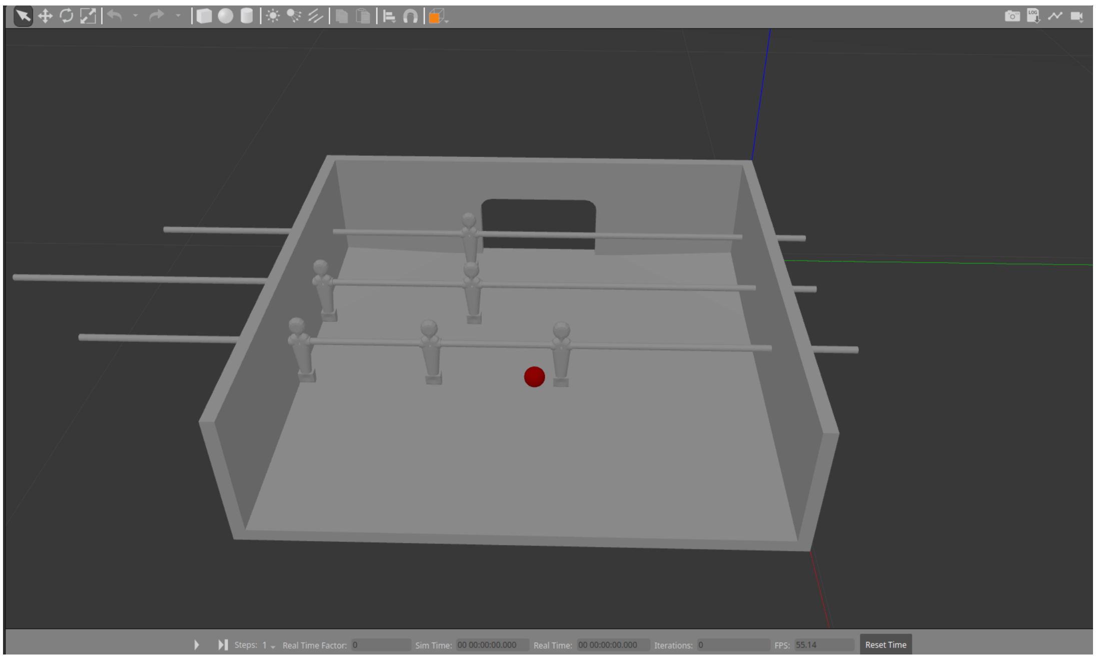

Figure 1: Simplified 3D model for ROS

2 Method description

The method used to tackle this problem was Proximal policy optimization (PPO) proposed by Schulman et al. (2017). The reward function was created to positively reward defending players for being in front of the ball and keeping the ball in the arena. The reward function also punished the defenders if they received a goal. The rotation had to be disabled as the action space was too big and due to non-converging over-fit experiments.

[1] Master-degree student of Applied Sciences and Informatics, field of study Cybernetics and Control Engineering with specialization Artificial intelligence and biocybernetics, e-mail: sieberm@students.zcu.cz

3 Results

The reward function over the period of learning can be seen in Table 1. The results from in silico training can be seen in Table 2. Finally results from real arena are in Table 3, the expert system was inspired by my previous work in Sieber (2020).

Model	Average reward per episode
Model after 500k time steps	51
Model after 1.5M time steps	95
Model after 2.5M time steps	105
Model after 4.5M time steps	124

Table 1: Average reward over time steps

Dataset	Success rate 500k steps	Success rate 4.5M steps	Random action
Training	69.4 %	81.2 %	65.6 %
Validation	49.4 %	61.2 %	55.3 %

Table 2: Success rate of defending on data sets

Model	Success rate of defense
Random	28%
RL model	62%
Expert system	70%

Table 3: Success rate on real arena with no rotation

The chosen approach had a major drawback, in the form of the need to create a shot dataset. This is due to the problem being not symmetrical, in a sense that defenders cannot score. This could be solved with an arena with two goals and equal players on both sides. This would eliminate the need to create dataset. The random action performed well in the simulation due to faster actuators and was surpassed in the real arena. The model in the later stages of training began to show the desired behavior corresponding to the reward function. This problem requires a longer training period and ideally automatic shot generation. Although the results are not as good as I wanted, this work provides the simulation environment for further experiments and a baseline method.

References

Schulman, John and Wolski, Filip and Dhariwal, Prafulla and Radford, Alec and Klimov, Oleg, arXiv preprint arXiv:1707.06347 (2017)

Sieber M,Detekce pohybu míčku pro mechatronický model stolního fotbalu, ZČU