Control of a robotic system based on image data in a simulation environment

Ing. Andrej Michalek **Supervisor: Ing. Peter Tarábek, PhD.**

Faculty of Management Science and Informatics
University of Žilina

FRI UNIZA

Motivation

Research in the field of autonomous driving is popular, especially in connection with driving vehicles in road traffic. The input to control systems is often location data, while the system has a map available to make decisions. However, an accurate map may not always be available, especially in an industrial environment where new obstacles may be added at random locations. In addition to navigating the environment, it is often necessary to identify an object and interact with it. An example can be the situation we dealt with in this diploma thesis. The forklift is supposed to identify the pallet, navigate to it and slide the skids into it without colliding into the obstacles in the environment.

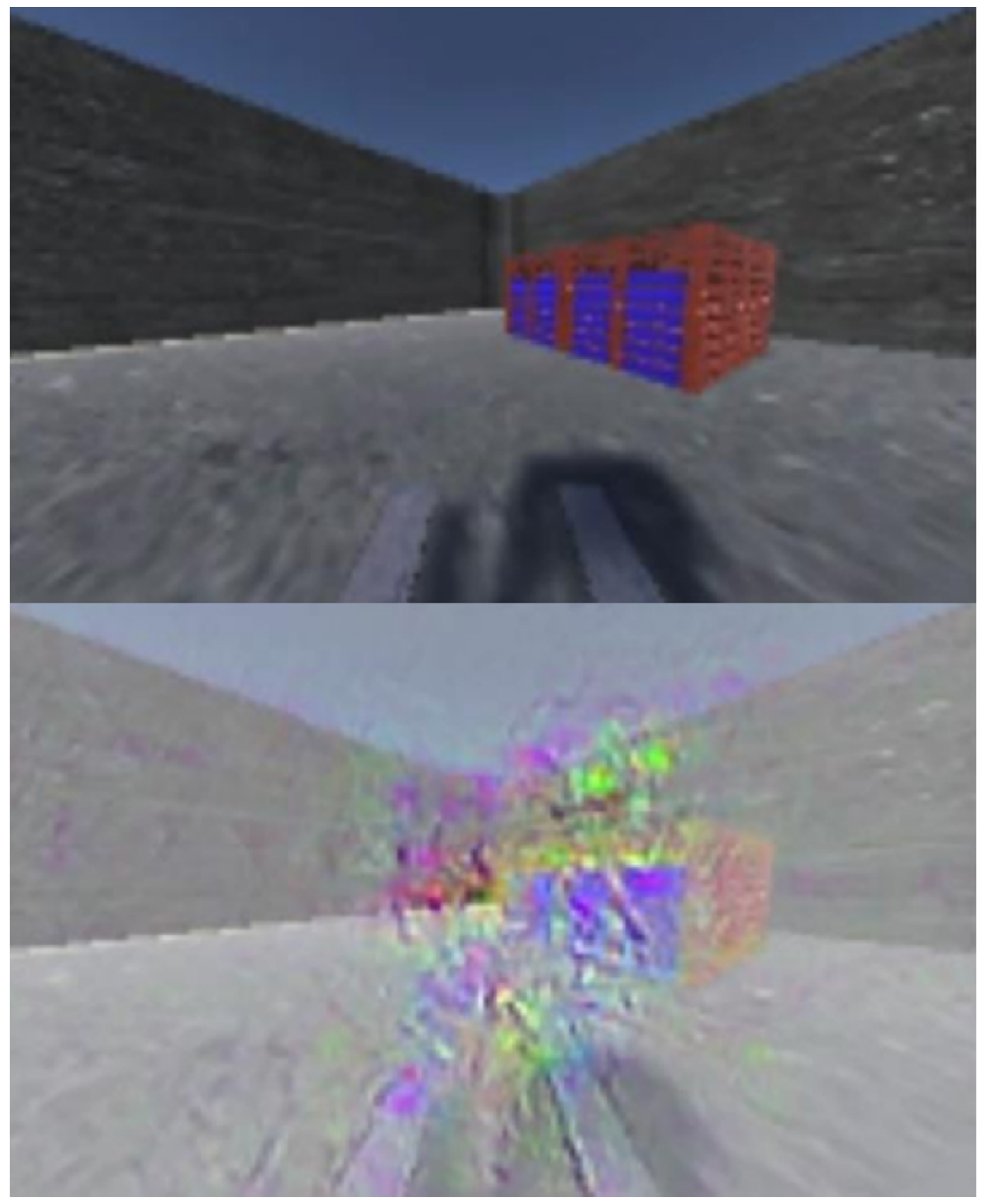
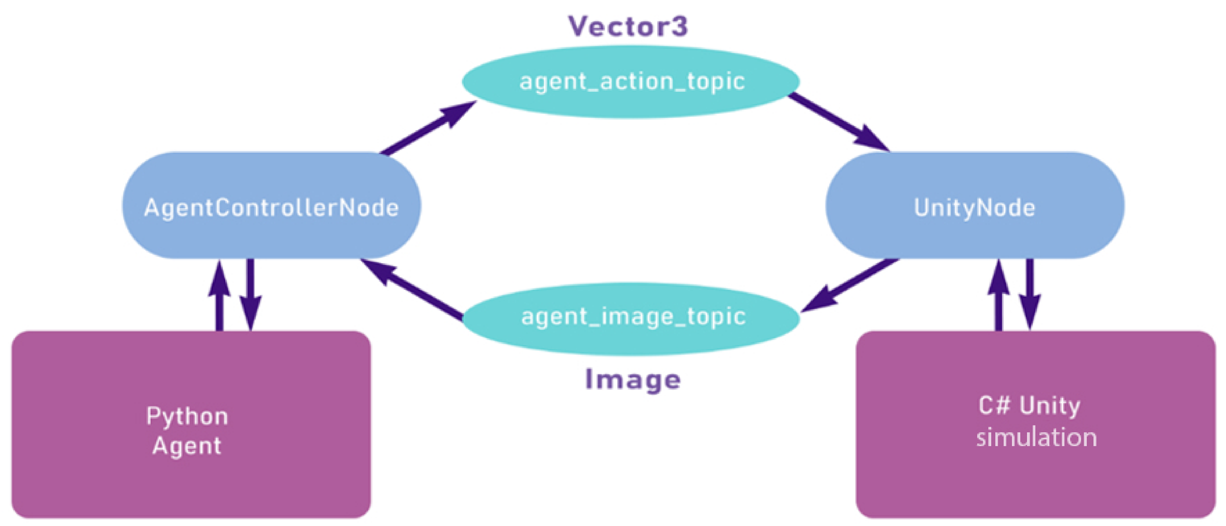


Image input for the agent's decision and visualization of the parts of the image that had the greatest impact on the decision

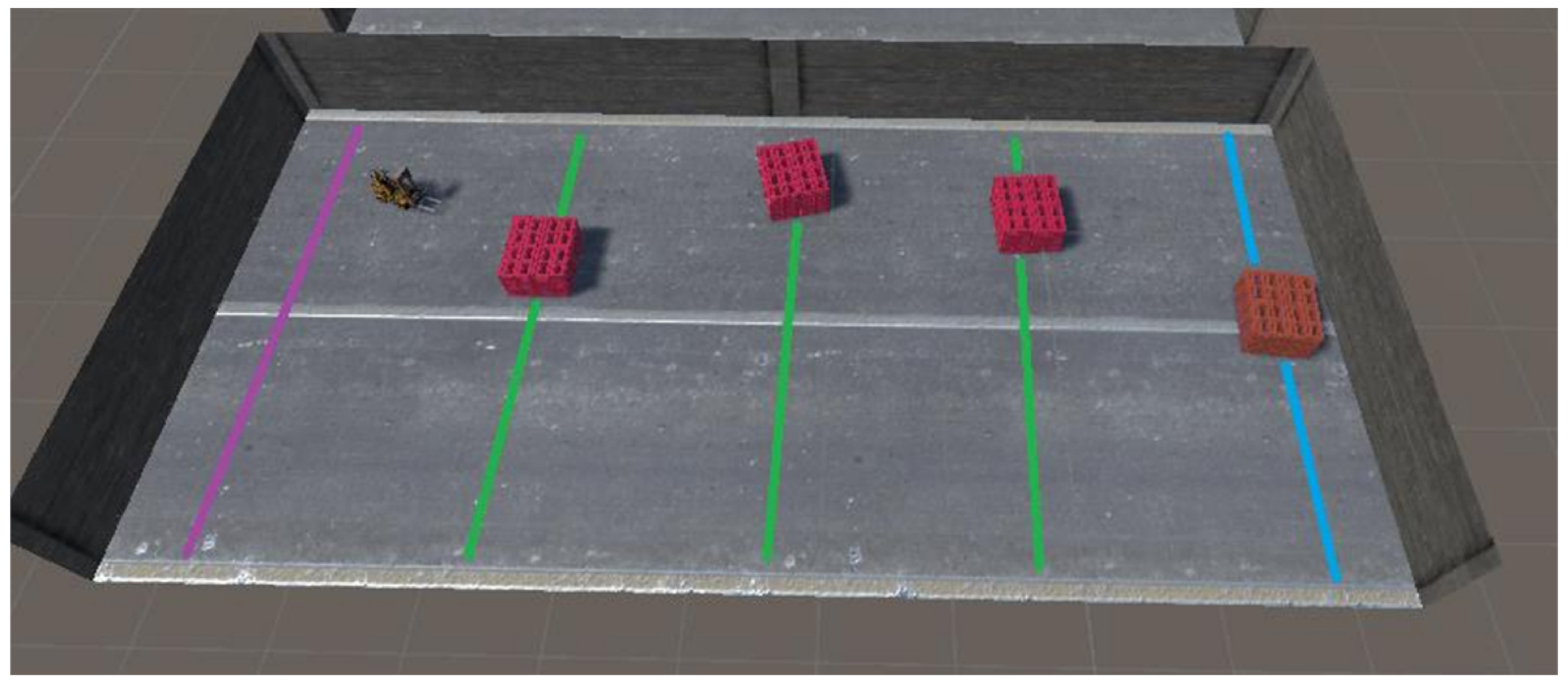
Agent and simulator

We used reinforcement learning as a learning tool for the agent. At the beginning, the agent performed random actions, the environment monitored his behavior and sent him feedback in the form of a reward. Based on the reward, the agent gradually improved his strategy.



Communication between the agent and the environment

In the learning process, the agent must perform a large number of attempts, so it would be impractical to train it in a real environment. Therefore, a simulator was created in the Unity engine. Communication has been ensured through ROS 2, so that multiple agent implementations can be exchanged in one environment and that the agents are ready for use on real vehicles without major changes. The environment enables training on multiple vehicles at the same time, which improves the use of parallel resources.

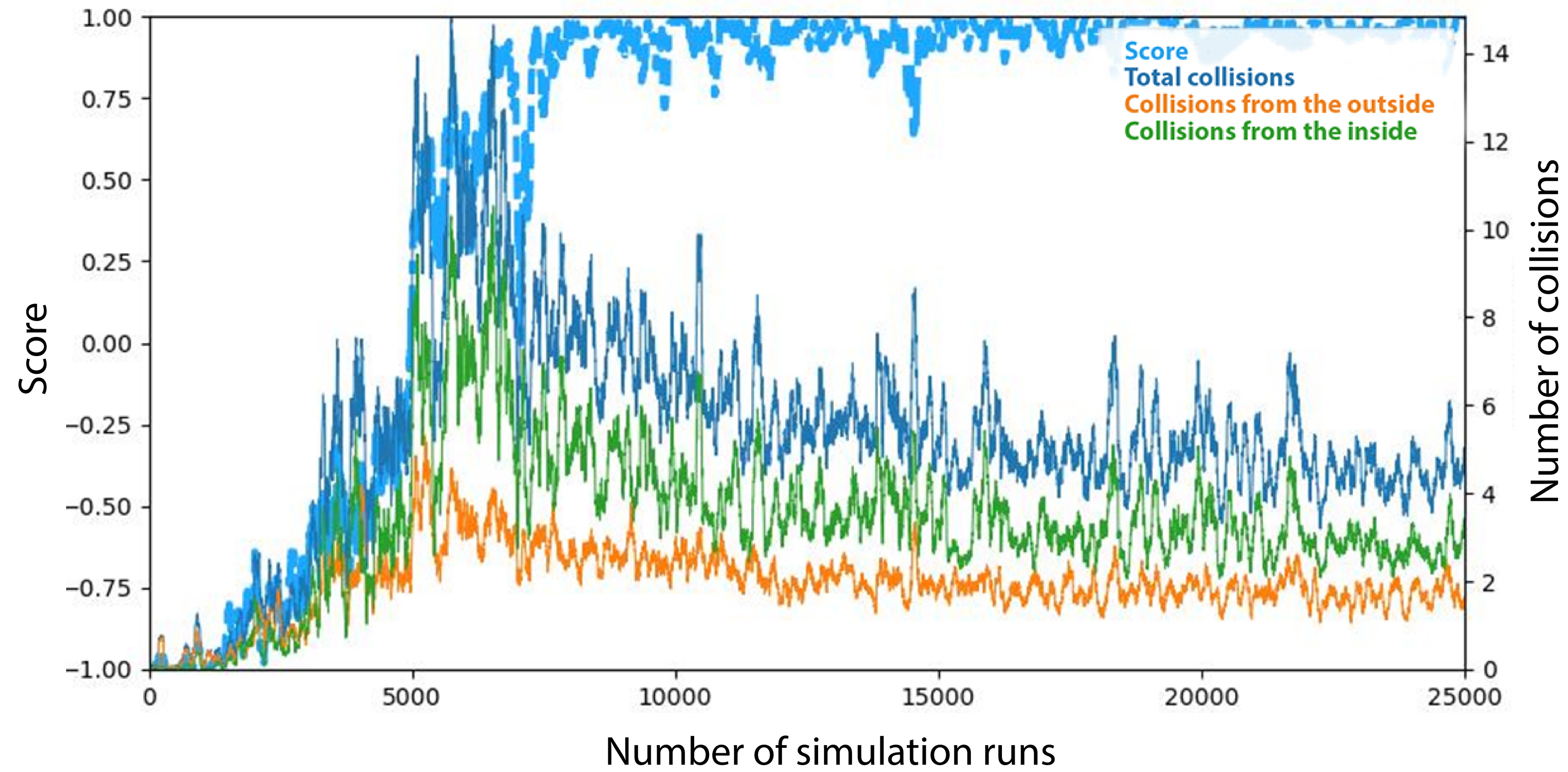


Environment sample - obstacles were generated randomly on green lines, forklift position on purple line and target palette on blue line

Experiments

In an effort to accelerate the training process and achieve more natural agent behavior, we conducted several experiments:

- Use of different types of agents - DQN, DDQN and PPO
- Test sensitivity to changing image resolution and state space expansion
- Using different approaches to a synchronization between agent and environment
- Testing generalisation abilities using a narrower palett, more complex environment and changing textures
- Sensitivity to changes in vehicle dynamics



Training of PPO agent using an image supplemented with a depth map. The agent in this case had a success rate of approximately 94%

Conclusion

As part of this thesis, it was possible to find a suitable agent, set its parameters and state and action space so that it could effectively solve the given task. We also looked at the vanilla gradient explainability method, which showed which information the agent uses in the decision-making process. Experiments have shown that the chosen agent can adapt well to changes in the environment through training, and thus it is assumed that it would be suitable for practical use as well.