UNIVERSITÄT
LEIPZIG

Master Thesis

Investigating the Feasibility of Training an Agent to Drive a Vehicle in a Simulated Environment Using Reinforcement Learning with Visual Camera Input

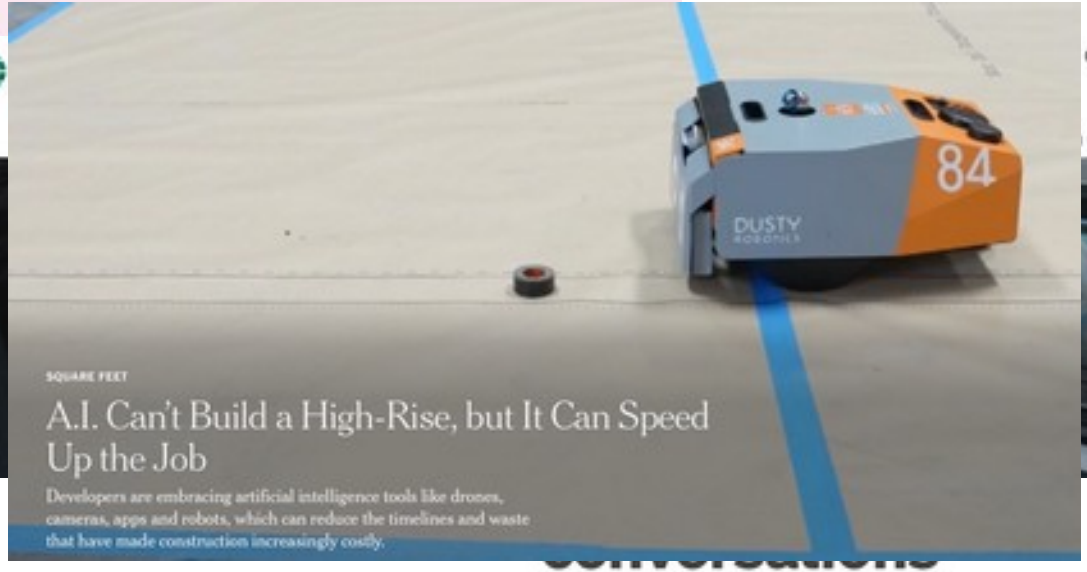
Leipzig, 08.09.2023

Maximilian Schaller

BLOG POST
RESEARCH

20 JUL 2016

Deep
Goog
Cooli



SQUARE FEET

A.I. Can't Build a High-Rise, but It Can Speed Up the Job

Developers are embracing artificial intelligence tools like drones, cameras, apps and robots, which can reduce the timelines and waste that have made construction increasingly costly.

Topics Te Ao Māori P

prises

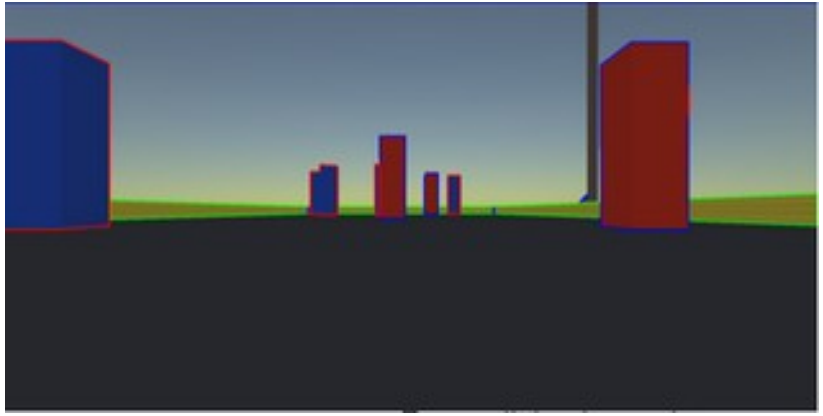
From **Morning Report**, 8:47 am on 7 December 2022 Share this

Master Thesis

PREVIOUS WORK

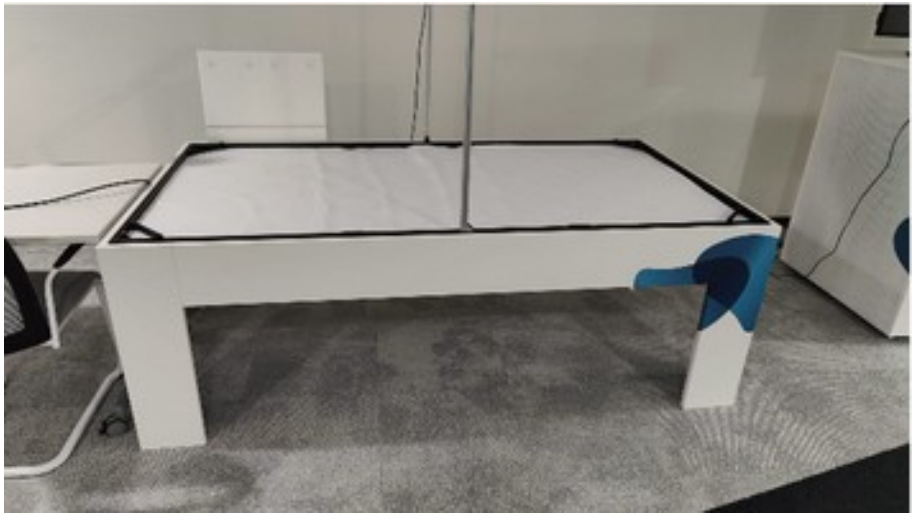
BACHELOR THESIS 1

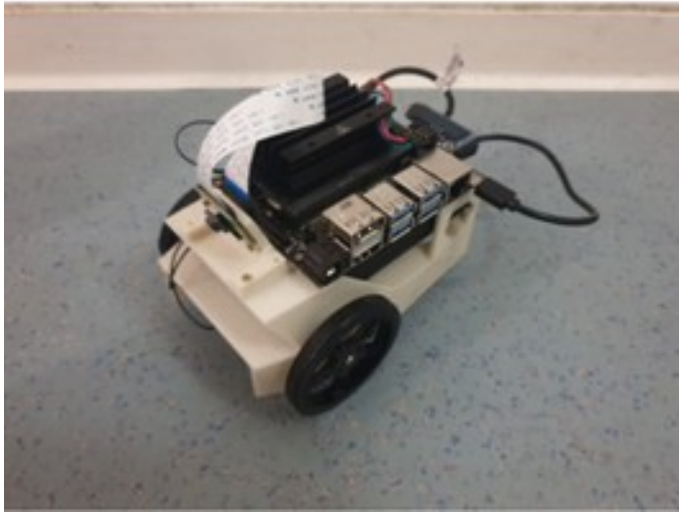
MODEL TRAINING OF A SIMULATED SELF-DRIVING VEHICLE USING AN EVOLUTION BASED NEURAL NETWORK APPROACH



BACHELOR THESIS 2

METHODS TO BRIDGE THE SIMULATION-TO-REALITY GAP

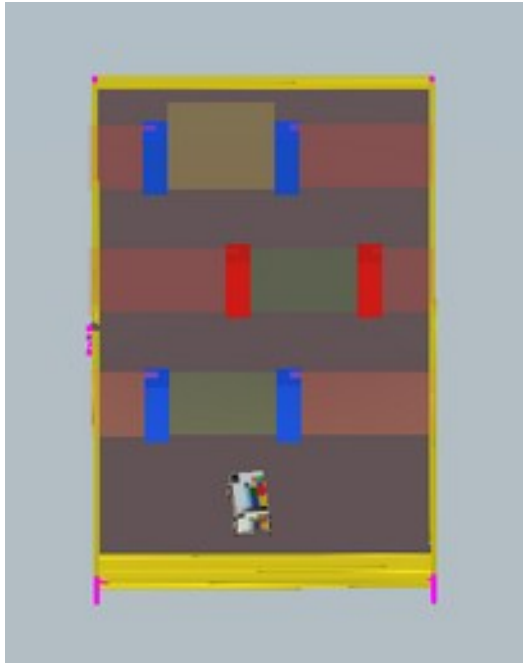




Master Thesis

RESEARCH QUESTIONS

GENERAL TASK



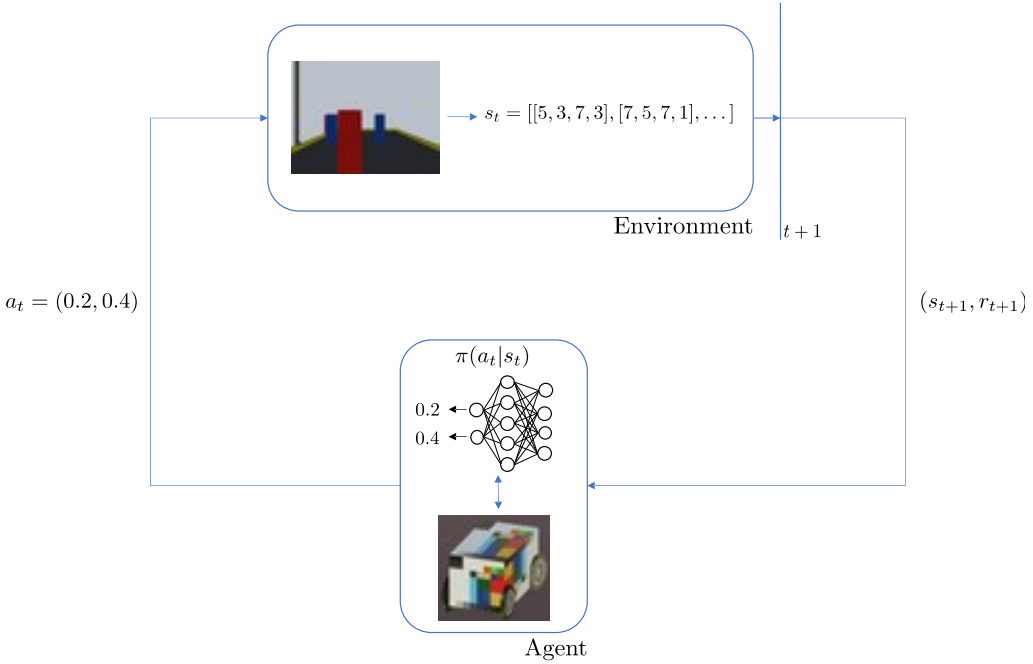
RESEARCH QUESTIONS

1. Is it possible to model the problem of training an autonomous driving agent in the context of reinforcement learning?
2. How can visual camera input be processed and utilized to achieve learning success?
3. How suitable is RL in solving the problem of training an agent to drive a vehicle and pass all goals?
4. How robust are the developed RL algorithms against varying external influences?

Master Thesis

BACKGROUND

REINFORCEMENT LEARNING



Agent

- Performs actions according to the policy

Policy

- Evaluates the state and proposes the best action

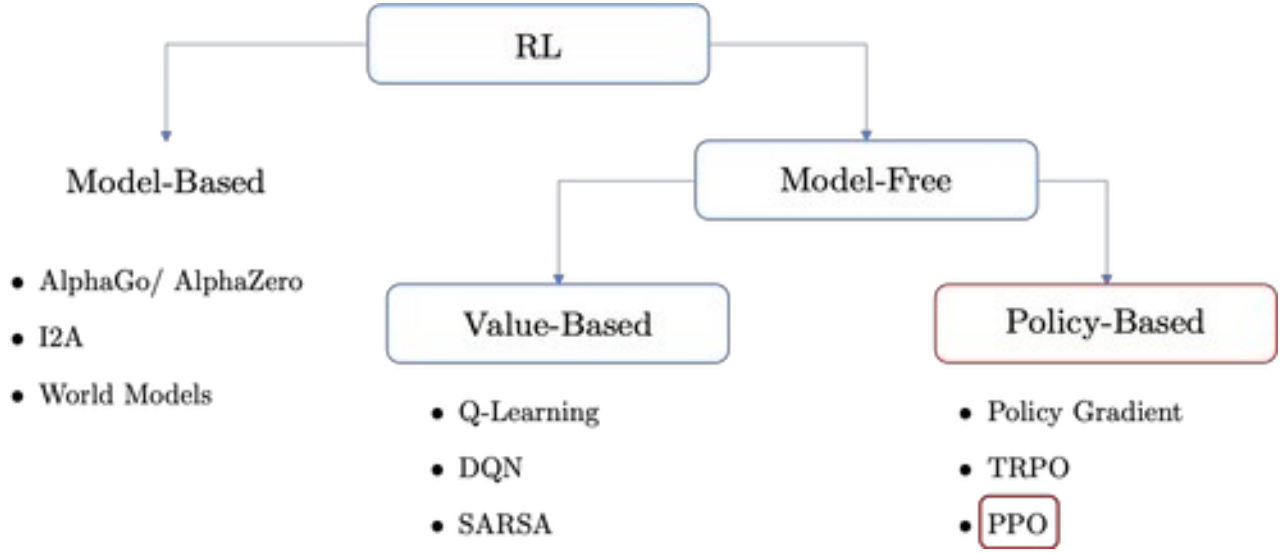
Environment

- Responds to an action wit a new state and a reward

Reward

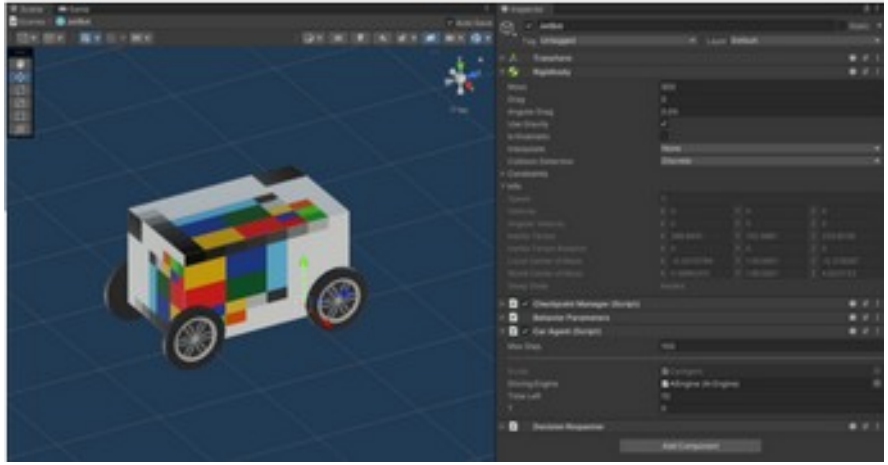
- Depicts the value of an action in a state

REINFORCEMENT LEARNING

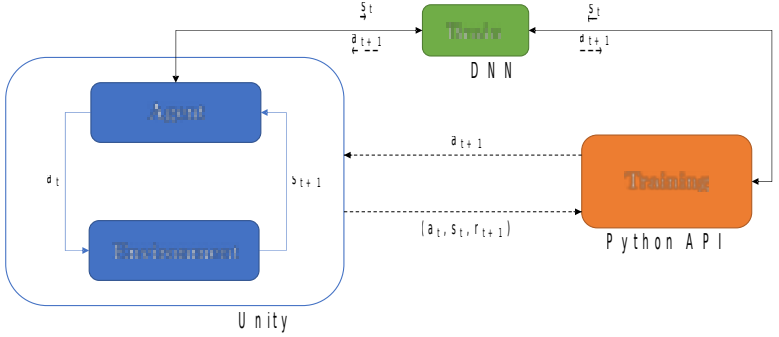


SIMULATION

UNITY - 3D DEVELOPMENT PLATFORM



ML-AGENTS FRAMEWORK

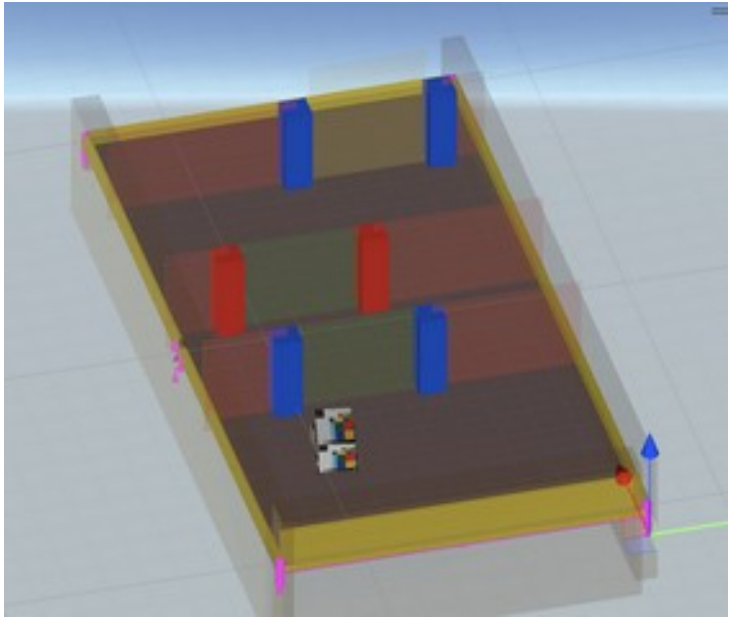


Master Thesis

IMPLEMENTATION

SIMULATION

ARENA



VEHICLE

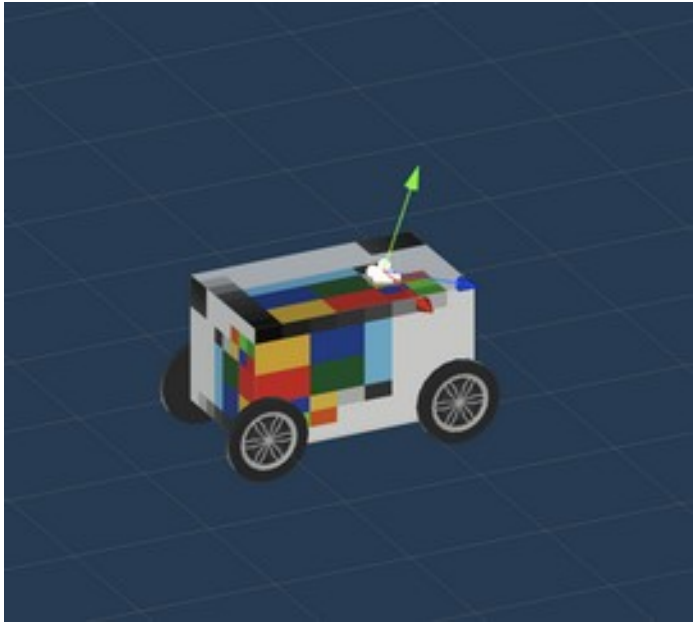
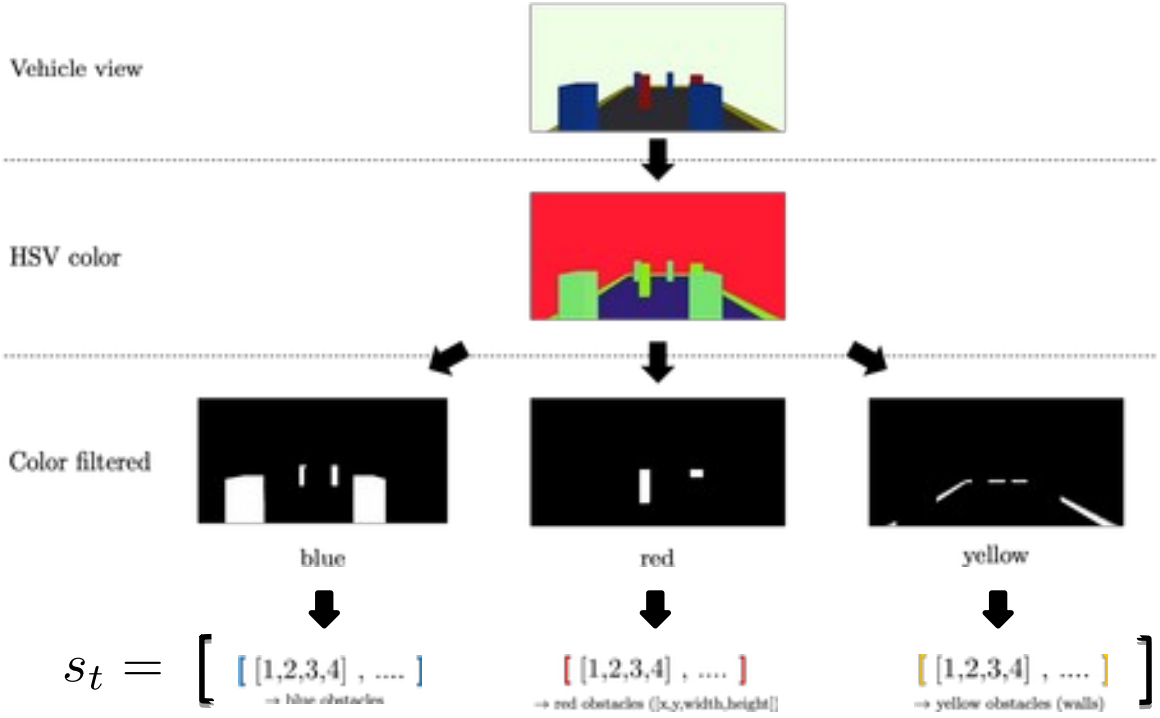


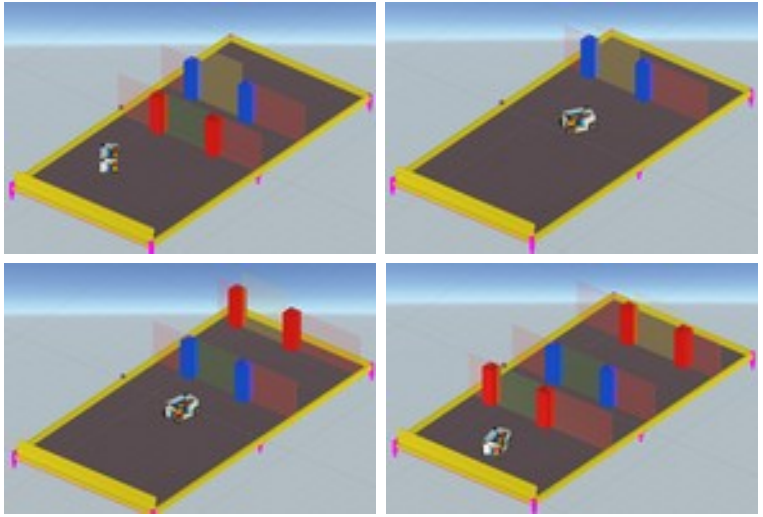
IMAGE PRE-PROCESSING



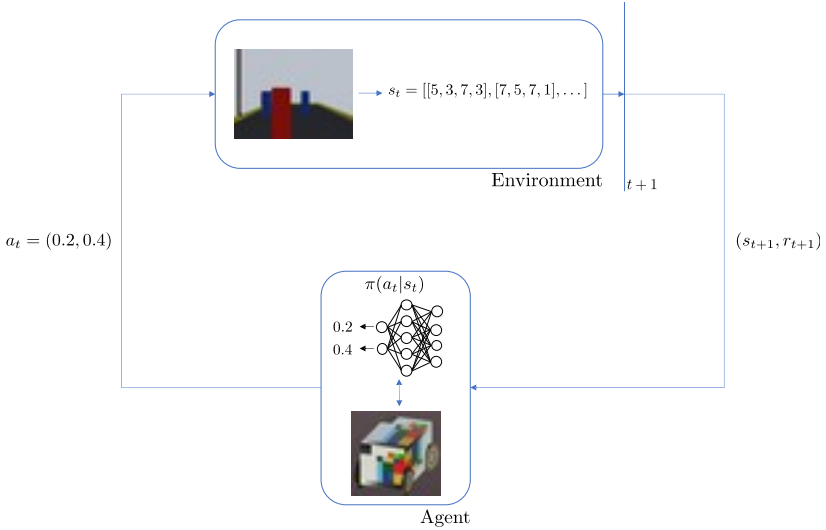
- reduction of 99.9997% compared to an 8-bit grey-scaled 1640x1232 picture

TRAINING PROCEDURE

INITIALIZATION

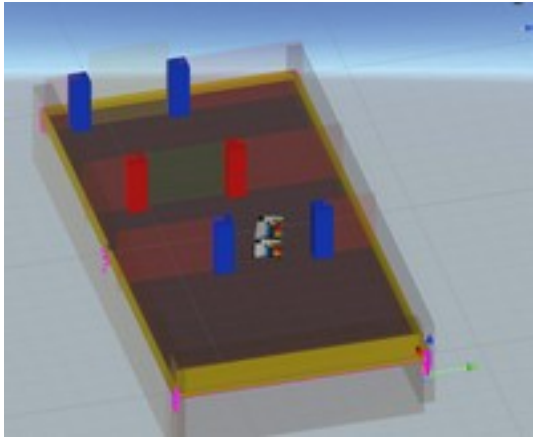


RL CYCLE

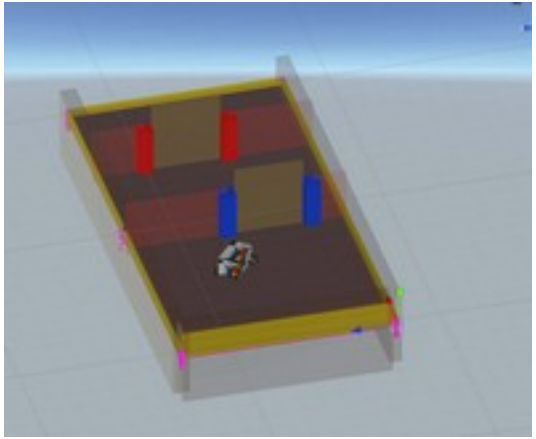


TRAINING VARIATIONS

FULL-MAP-TRAINING (FMT)



SINGLE-GOAL-TRAINING (SGT)



PROXIMAL POLICY OPTIMIZATION ALGORITHMS

STATE-ACTION-REWARD MODELS

Action: - accelerate left motor (a_1) and right motor (a_2) $a_{1|2}$ (-1,1)

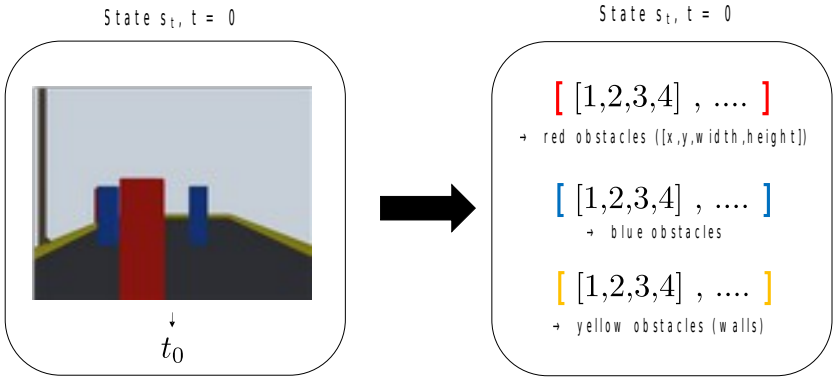
Reward: - distinct for each training variation R_{SGT} & R_{FMT}

$$R_{SGT}(s_t, a_t) = \begin{cases} (v/10) * \Delta T, & \text{for } \forall t \\ 1, & \text{for } \textit{Goal passed} \\ -1, & \text{for } \textit{Goal missed} \\ -1, & \text{for } \textit{Hit post} \\ -1, & \text{for } \textit{Hit wall} \\ -1, & \text{for } \textit{Out of time} \end{cases}$$

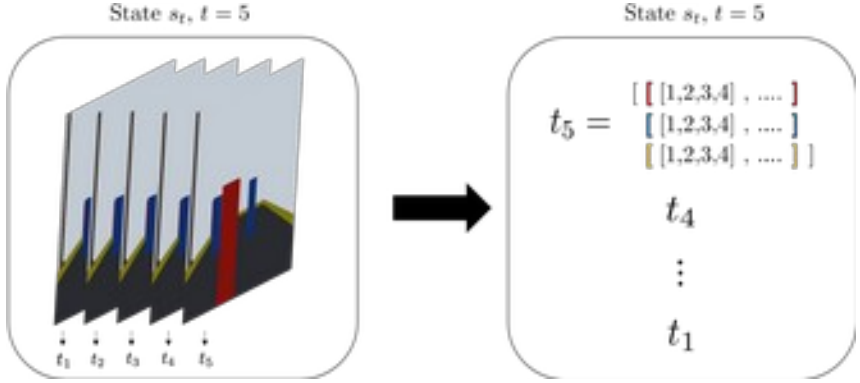
$$R_{FMT}(s_t, a_t) = \begin{cases} (v/10) * \Delta T, & \text{for } \forall t \\ 1, & \text{for } \textit{Goal passed} \\ 100, & \text{for } \textit{Map completed} \\ -1, & \text{for } \textit{Goal missed} \\ -1, & \text{for } \textit{Hit post} \\ -1, & \text{for } \textit{Hit wall} \\ -1, & \text{for } \textit{Out of time} \end{cases}$$

TRAINING VARIATIONS

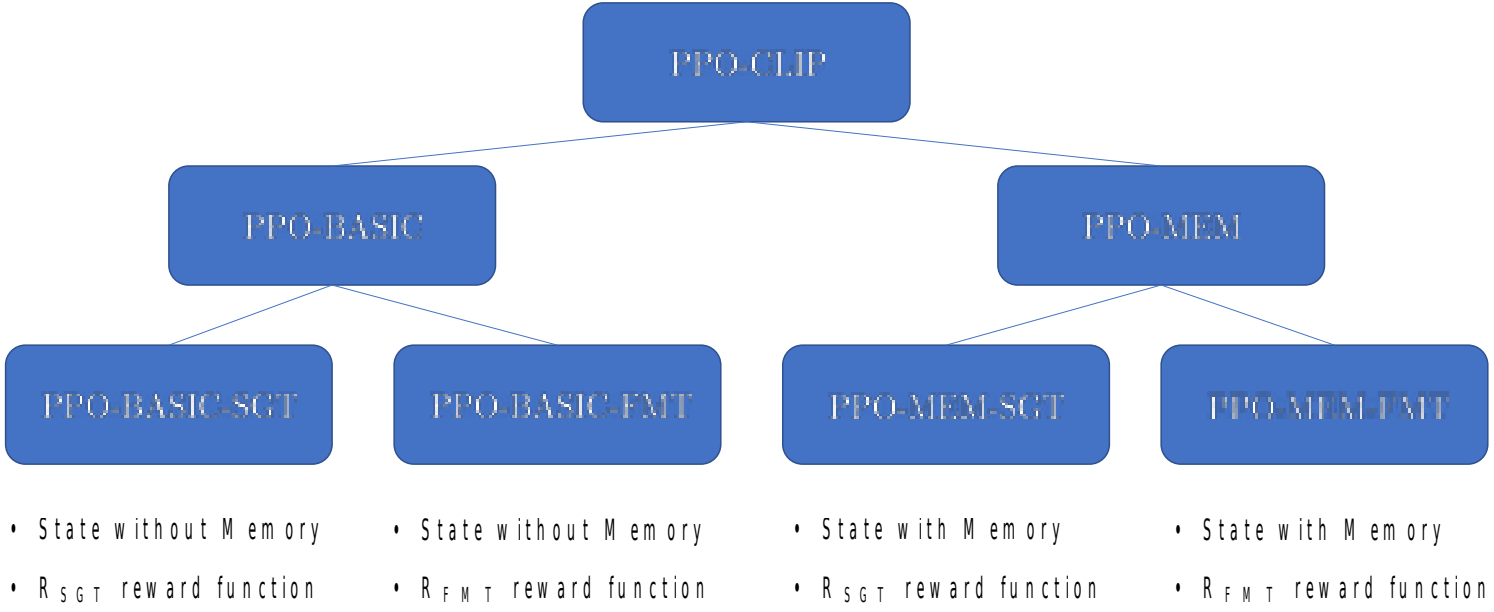
STATE WITHOUT MEMORY (BASIC)



STATE WITH MEMORY (MEM)

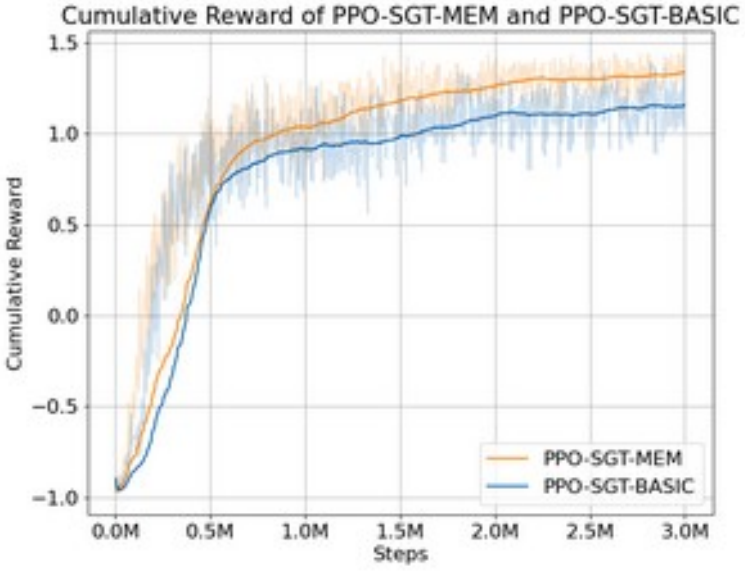


PPO ALGORITHMS

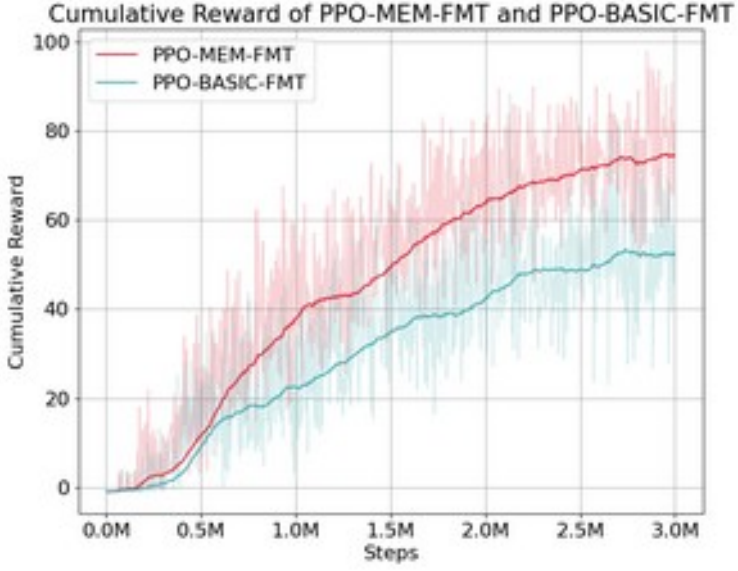


TRAINING PERFORMANCE

SGT APPROACHES



FMT APPROACHES

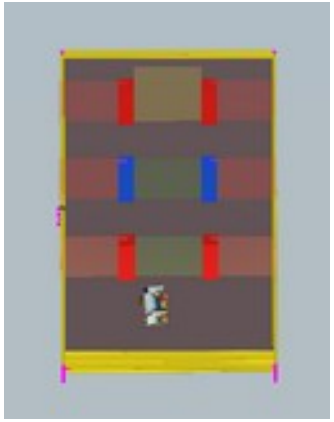


Master Thesis

EXPERIMENTATION

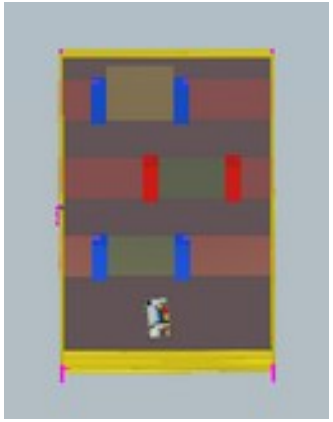
MAP DESIGN

EASY



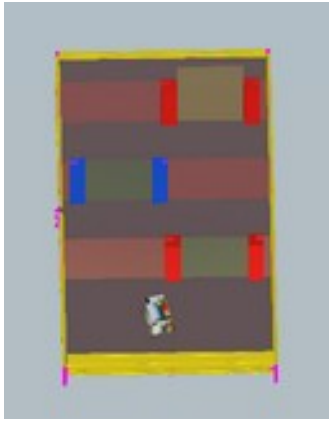
- 2 maps (starting with blue or red)

MEDIUM



- 4 maps each: first goal blue or red and left (2) + blue or red and right side (2)

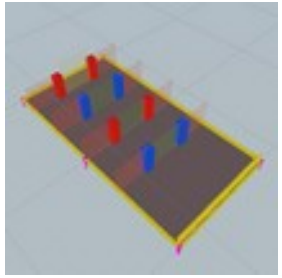
HARD



- 10 different map configurations = 2 easy + 4 medium + 4 hard
- 100 runs on every configuration in every experiment = 1000 runs per experiment

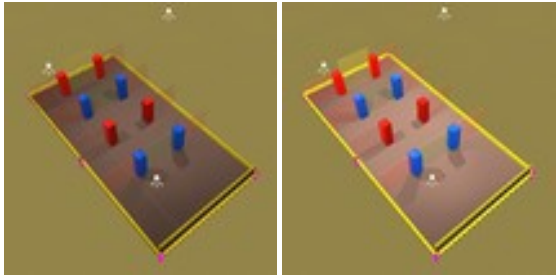
EXPERIMENTS

1 – OPTIMAL CONDITIONS



- Ambient light, no shadow
- Motor power as in training

2 – DIFF. LIGHT SETTINGS

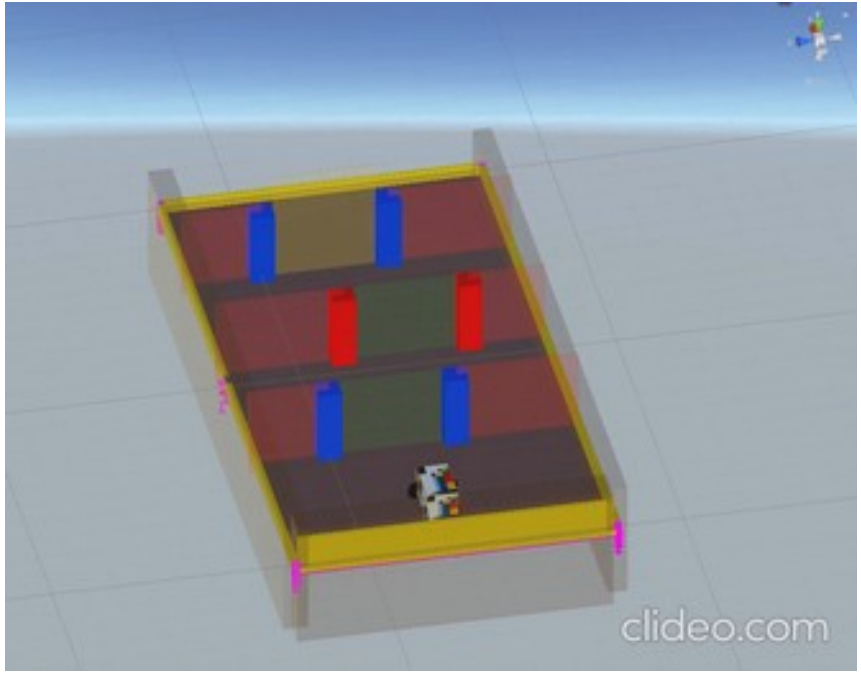


- 6 spot lights
- Dark light setting (left) and bright light setting (right)
- Lights causing shadow
- Motor power as in training

3 – VARYING MOTOR POWER

- Ambient light as in training
- +20% motor power
- -20% motor power

EXPERIMENT 1 - EXAMPLE

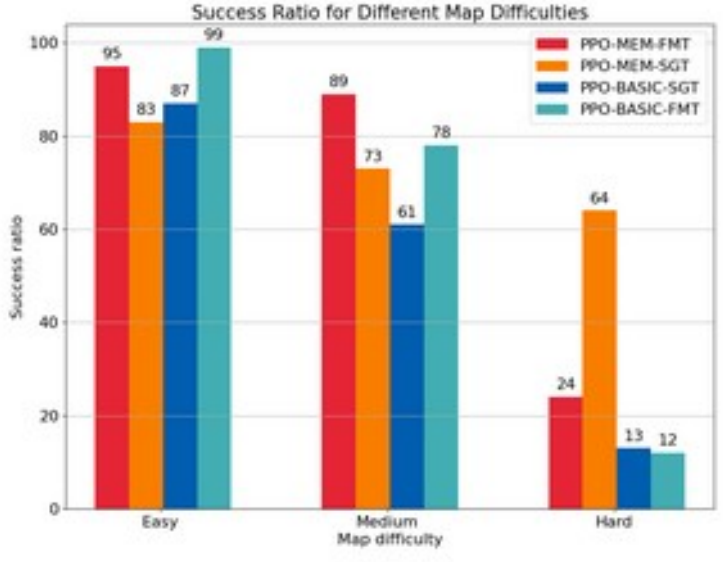


Master Thesis

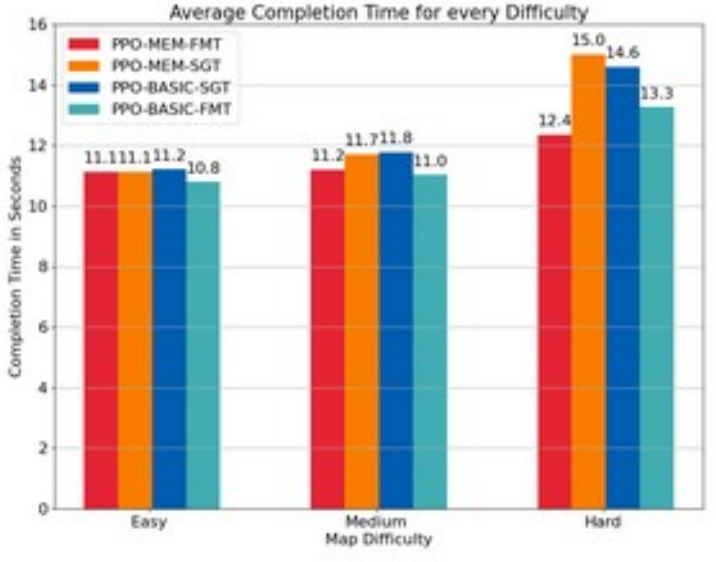
EVALUATION

EXPERIMENT 1

SUCCESS RATIO PER MAP



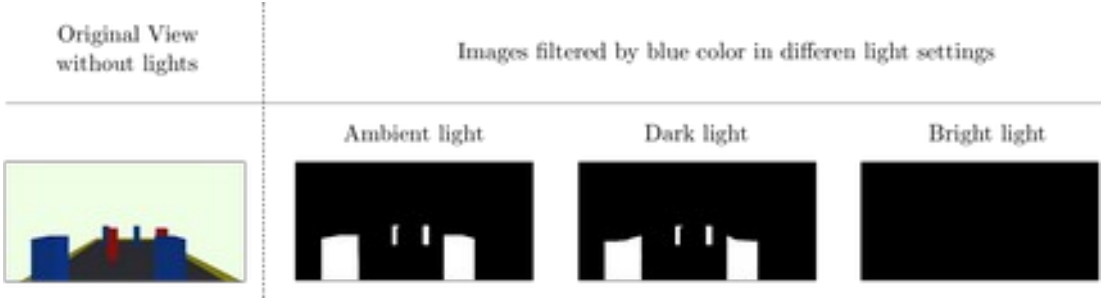
AVG TIME TO COMPLETE A MAP



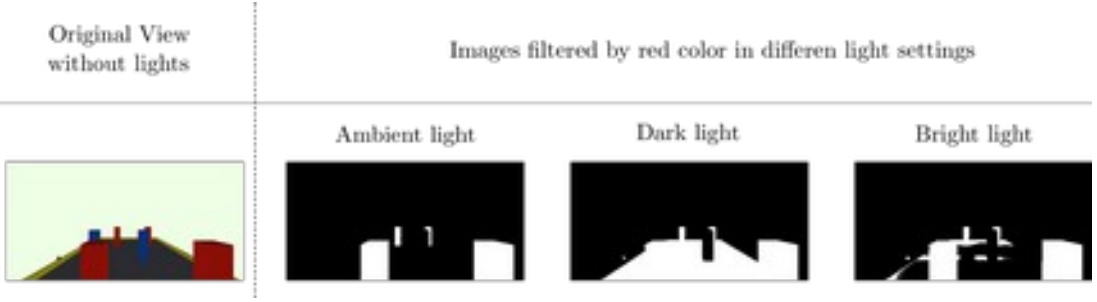
EXPERIMENT 2

NOT STABLE IN DIFFERENT LIGHT SETTINGS DUE TO IMAGE PRE-PROCESSING

Blue obstacles

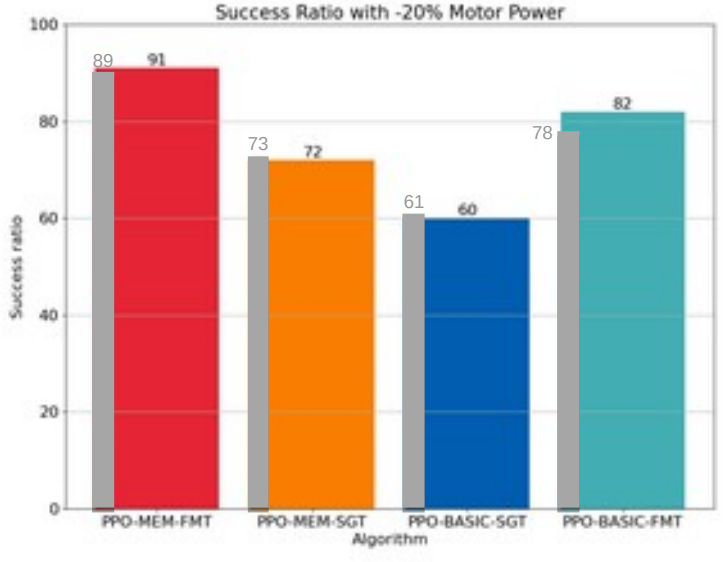


Red obstacles

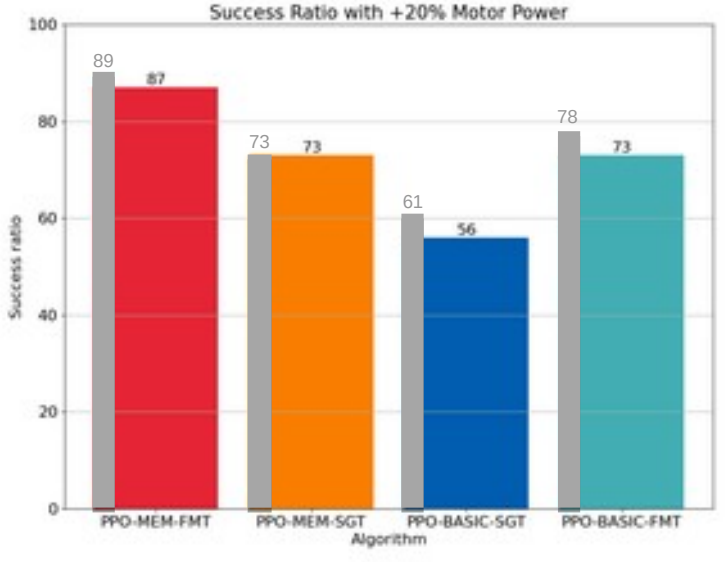


EXPERIMENT 3

SUCCESS RATIO -20% ENGINE POWER



SUCCESS RATIO +20% ENGINE POWER



Master Thesis

CONCLUSION & FUTURE WORK

RESEARCH QUESTIONS

1. Is it possible to model the problem of training an autonomous driving agent in the context of reinforcement learning? 
2. How can visual camera input be processed and utilized to achieve learning success? 
3. How suitable is RL in solving the problem of training an agent to drive a vehicle and pass all goals? 
4. How robust are the developed RL algorithms against varying external influences? 

FUTURE WORK

- more training time & more seeking for better hyperparameter
- Add input sources (sensors, more cameras, …)
- Train in varying influences such as light settings or motor power
- Improve the pre-processing, e.g. by using CNNs
- Improve the memory, e.g. by utilizing LSTMs

UNIVERSITÄT
LEIPZIG

THANK YOU!

Maximilian Schaller
Scads.AI
Faculty of Mathematics & Computer Science
Database Department

ms80xuxu@uni-leipzig.de
www.uni-leipzig.de