TECHNISCHE
UNIVERSITÄT
WIEN

Master Thesis Proposal

Attention Based Neural Network for Autonomous Driving Agents

by **Stefan Ulmer**, MSc (WU) BSc

Matr. number: 11777750, Master Program Computer Engineering 066.938

E-Mail: e11777750@student.tuwien.ac.at

TU Vienna, May 9, 2022

Supervisor: Head Univ.Prof. Dipl.-Ing. Dr.rer.nat. **Radu Grosu**

Co-Supervisor: Univ.Ass. Dipl.-Ing. / BSc **Andreas Brandstätter**

Contents

1 Motivation & Problem Statement

Autonomous driving is a very actual and important topic that reforms the car industry. The usage of an increasing number of sensors and actuators as well as deep neural networks for a car's automatic action are necessary to provide safe and stable driving. Implementing **single agents** with good decision-making capabilities is hard, but required to reach this. Classical approaches cannot capture all possible upcoming events and incidents, a promising solution is the use of machine-learning.

Different types of single agents are commonly used for the implementation of autonomous driving algorithms. Starting from simple reflex to more complex goal-based agents they all have to consider the current state of the agent's environment, in order to take an appropriate action, possibly in a probabilistic fashion. From this state the agent can take an action that leads to the following state [8].

Even tough reinforcement learning is an emerging area in autonomous driving, the applications that are commercially and successfully used by now are manageable. The difficulties for this task arise from the scarce availability of datasets and specific literature [4]. Recent studies include attention as a mechanism to interpret and improve algorithms used for deep learning [3].

In virtual reality, attention already shows improvements for existing networks [6]. In correspondence to that, it is interesting to consider humans **attention to decide a vehicle's control** for a specific state. In this context, it is hard to synthesize models and algorithms for driving behaviour without compromising safety. Since humans are able to cope with complex driving scenarios, a **human-inspired attention mechanism could improve autonomous driving**.

2 Aim of the Work

The aim of this thesis is the development of a stable and reliable agent based on an attention neural network. The first step is the hardware setup which includes eye-tracking technology for the creation of an appropriate dataset. By using the eye-tracking system the human's behaviour will be recorded and then the attention will be emulated by training a neural network model. By leveraging the information extracted by the attention model, a novel and more efficient training of the policy model should be developed. The **attention neural network** should be able to predict the attention focus on a specific frame input from the real car's camera.

Furthermore, an **agent neural network** has to be trained with the attention dataset to learn the necessary car actions for different driving scenarios. This agent model has to take actions based on the input frame with the predicted attention focus from the attention neural network. For validating the resulting single-agent controller, simulations on model predictions will be performed followed by deployment of the model on a real F1/10 racecar.

Existing literature about this novel attention mechanism supports the aim of this thesis: The implementation of an algorithm that is able to autonomously operate on a F1Tenth track [1].

At the end of this thesis, the **following questions** shall be answered:

- How much is a neural network's training- and prediction-performance improved by the attention mechanism?
- What differences are observable and measurable between the simulation and real car performance?
- What is the robustness of the trained agent to domain shift (e.g. different track)?

3 Methodological Approach

The effort of this thesis work can be split up to multiple work packages. Starting from the literature review, machine learning techniques have to be used for the training phase of a proper attention- and agent-model. After checking the model outcomes in the simulation environment, extensive analysis and evaluation will be done on the real F1/10 racecar.

3.1 Literature Review

For getting the necessary knowledge about autonomous driving and deep learning a review of the existing literature has to be done beforehand. Furthermore, based on this knowledge the aim is to get a deeper understanding of the F1Tenth-framework for the real car and the simulation.

3.2 Modelling

The machine learning process of the models for attention- and driving-prediction operates on the created eye-tracking attention dataset. By training a neural network on this human's attention dataset, a model has to be developed that predicts the attention focus for a specific frame input. For autonomously driving on a track, an agent neural network has to be trained to predict the car's action based on an attention focused frame. These models will be used for controlling the car within the driving environment.

3.3 Simulation

The simulations of the car's behaviour will be done with the mentioned F1Tenth-framework. Therefore the specific model predictions will be integrated to ensure reasonable outcomes in simulation. This has to be done for making sure that the models work properly before the deployment on the real car.

3.4 Analysis

When simulation and models have reached the desired level of accuracy, it can be considered to drive the real car on a track. The analysis of the performance on a racecar will drive the additioned refinements and possible iterations in simulation to ensure safety on the real track.

3.5 Evaluation

Having setup the complete pipeline from training to deployment on the real car, the approach has to be evaluated compared to the state of the art. The evaluation phase will consider the research questions and will compare the performance between the simulation and the real car.

4 State of the Art

Current approaches for synthesizing single-agent policies are based on deep reinforcement learning algorithms. These techniques have been proned to be effective to produce collision-free behaviors [9] [7]. Autonomous driving often includes additional subsystems that has to be integrated: Localization, Perception, Planning

[2]. The increased usage of sensors (e.g. LiDAR) and the combination of computer vision and hardware acceleration pushed the improvements of autonomous driving, whereas the overall goals would be reliable results in real-time. Recent systems are in continuously development for real traffic environments as current computing systems are still not reliable and robust enough for autonomous driving on level 4 or 5 [5] [10].

However, Wang and Yang [12] show that unsupervised models present sub-optimal performance. They included attention to their model, such that the neural network model uses supervised learning with a recurrent neural network. Furthermore, Vaswani et al. [11] stated that the best performing models integrate an attention mechanism and came up with a new architecture - the Transformer. They proved that in translation tasks, the models need less training time for achieving adequate performance. Makrigiorgos et al. [6] shows that in simulations and virtual reality these models improve the accuracy.

Nevertheless the existing literature about attention and its usage within a neural network is manageable. Additionally, there is a lack of information about testing such model systems on real cars. Hence, the result of this thesis will show a real F1Tenth-Car driving performance based on model predictions that include an attention mechanism.

5 Structure of the Work

1. Abstract
2. Introduction
    1.1 Motivation
    1.2 State of the Art
    1.3 Methodological Approach
3. Attention Neural Network
    2.1 Eye-Tracking Technology
    2.2 Attention Mechanism
    2.3 Model Creation
4. Agent Based Behaviour
    3.1 Decision Making
    3.2 Agent Implementation
5. Analysis
    4.1 Attention Model Outcome

4.2 Simulation vs. Real Hardware Car
4.3 Answering Research Questions
6. Conclusion and Future Work

6 Topic Match with Computer Engineering

This Master-Thesis contributes to the existing state of the art and uses an attention neural network based on eye-tracking technology. The outcoming model is used for autonomous driving on real hardware, i.e. on a F1Tenth-Car.

Therefore a lot of topics of the Computer Engineering studies will be covered during this thesis. Especially the following courses have an important role:

- 182.763 - Stochastic Foundations of Cyber-Physical Systems
- 191.119 - Autonomous Racing Cars
- 182.762 - Project Computer Engineering (IROS Race Competition)
- 376.054 - Machine Vision and Cognitive Robotics
- 183.660 - Mobile Robotics
- 182.753 - Internet of Things

References

[1] Varundev Suresh Babu and Madhur Behl. "f1tenth. dev-An Open-source ROS based F1/10 Autonomous Racing Simulator". In: *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE).* IEEE. 2020, pp. 1614–1620.

[2] Mrinal R. Bachute and Javed M. Subhedar. "Autonomous Driving Architectures: Insights of Machine Learning and Deep Learning Algorithms". In: *Machine Learning with Applications* 6 (2021), p. 100164. ISSN: 2666-8270. DOI: https://doi.org/10.1016/j.mlwa.2021.100164. URL: https://www.sciencedirect.com/science/article/pii/S2666827021000827.

[3] Chen Chen et al. *Combination of deep neural network with attention mechanism enhances the explainability of protein contact prediction.* 2020.

[4] Bangalore Ravi Kiran et al. "Deep Reinforcement Learning for Autonomous Driving: A Survey". In: *CoRR* abs/2002.00444 (2020). arXiv: 2002.00444. URL: https://arxiv.org/abs/2002.00444.

[5] Liangkai Liu et al. "Computing Systems for Autonomous Driving: State-of-the-Art and Challenges". In: *CoRR* abs/2009.14349 (2020). arXiv: 2009.14349. URL: https://arxiv.org/abs/2009.14349.

[6] Alexander Makrigiorgos et al. "Human Visual Attention Prediction Boosts Learning & Performance of Autonomous Driving Agents". In: *CoRR* abs/1909.05003 (2019). arXiv: 1909.05003. URL: http://arxiv.org/abs/1909.05003.

[7] Arash Mohammadhasani et al. "Reinforcement Learning Based Safe Decision Making for Highway Autonomous Driving". In: *CoRR* abs/2105.06517 (2021). arXiv: 2105.06517. URL: https://arxiv.org/abs/2105.06517.

[8] Gonçalo Neto. *From Single-Agent to Multi-Agent Reinforcement Learning: Foundational Concepts and Methods Learning Theory Course*. 2005.

[9] Aizaz Sharif and Dusica Marijan. *Evaluating the Robustness of Deep Reinforcement Learning for Autonomous and Adversarial Policies in a Multi-agent Urban Driving Environment*. 2021. arXiv: 2112.11947 [cs.AI].

[10] Aizaz Sharif and Dusica Marijan. *Evaluating the Robustness of Deep Reinforcement Learning for Autonomous and Adversarial Policies in a Multi-agent Urban Driving Environment*. 2021. arXiv: 2112.11947 [cs.AI].

[11] Ashish Vaswani et al. "Attention is All You Need". In: 2017. URL: https://arxiv.org/pdf/1706.03762.pdf.

[12] Xinyi Wang and Yi Yang. "Neural Topic Model with Attention for Supervised Learning". In: *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*. Ed. by Silvia Chiappa and Roberto Calandra. Vol. 108. Proceedings of Machine Learning Research. PMLR, Aug. 2020, pp. 1147–1156. URL: https://proceedings.mlr.press/v108/wang20c.html.