Meta-Learning based Autonomous Driving Agents

Aakash Deshmane, Harsh Mittal, Humayun Akhtar, Shaunak Kolhe
Texas A&M University
College Station, TX
{deshmaneaakash, harshmittal27, humayun.akhtar, kolheshaunak}@tamu.edu

Abstract

Autonomous waypoint following is an established test case for autonomous driving scenarios, with implementations simple enough to be executed by standard proportional controllers, to higher order Deep Learning based controllers. In this project we aim to explore the design and implications of using a PID controller, a naive RL agent, and a Meta-Learning enabled RL Agent for a driving scenario with adverse conditions. The training and testing would be done in the CARLA simulation environment with variation in terrain, road characteristics.

1 Objective

In this research project, we aim to explore the design and implications of using Reinforcement Learning (RL) based agents for autonomous waypoint following in adverse weather conditions and varying terrain. The focus of the project will be to compare and analyze the performance of a fixed PID Controller, a naive RL agent and a Meta-Learning enabled RL agent Finn et al. [2017], Clavera et al. [2018]. The study will be conducted in the CARLA simulation environment to provide a controlled and reproducible environment for experimentation.

The naive RL agent will serve as a baseline for the comparison, implementing a standard Q-Learning algorithm. The fixed PID controller would add value as a conventional approach; while the Meta-Learning enabled RL agent will leverage Meta-Learning to quickly adapt to new situations, allowing for faster and more efficient training.

The results of the study will provide insights into the performance and capabilities of each of the RL agents, and the implications of using each for autonomous waypoint following in adverse conditions. Additionally, the results will contribute to the advancement of autonomous driving research by exploring the use of RL for controlling autonomous vehicles in challenging environments.

In this research project, we are not addressing a purely theoretical question. Our focus is on exploring the design and implications of using Reinforcement Learning (RL) based agents for autonomous waypoint following in adverse weather conditions and varying terrain and comparing it to the General Way point control and a naive RL controlÓscar Pérez-Gil et al. [2022]. Our approach is to train and evaluate three different control algorithms, including a naive RL agent, an general Way point follower, and a Meta-Learning enabled RL agent in the CARLA simulation environment.

The results of our study will be clearly demonstrated through a variety of visual aids, including plots, graphs, and animations. These visualizations will provide a clear understanding of the behavior and performance of each of the three agents, and will be used to compare and analyze the performance of each agent. Our results will be objectively evaluated using a set of performance metrics and statistical tests, providing a clear and impartial evaluation of the capabilities and limitations of each agent for autonomous way point following in adverse weather conditions and challenging terrain.

Preprint. Under review.

2 Literature Review

In our research project, we aim to leverage the concept of Meta Reinforcement Learning (MRL) Arnold et al. [2020] to develop an adaptive control system for autonomous vehicles to follow waypoints in navigating unseen terrain. Our approach is inspired by the existing work on Rapid Motor Adaptation (RMA) algorithm for legged robots as presented in Kumar et al. [2021].

The RMA algorithm consists of two components: a base policy and an adaptation module. The base policy is trained through reinforcement learning in simulation, incorporating information about the environment configuration such as friction and payload. The environment configuration is first encoded into a latent feature space using an encoder network, which is then fed into the base policy along with the current state and previous action. The adaptation module, on the other hand, is trained to estimate the latent feature space from the robot's recent state and action history and is used in real-time to generate the environmental information during deployment.

In our research project, we plan to enhance the existing implementation of the RMA algorithm in several ways. This includes improving the training process and using a more diverse set of environments for training. Our ultimate goal is to evaluate the performance of the algorithm on various real-world scenarios, such as uneven ground, slippery surfaces, and rough terrain, to ensure its robustness and adaptability for autonomous vehicle control.

3 Simulation Environments

For all stages of training and testing, we aim to the CARLA Dosovitskiy et al. [2017] environment with multiple scenarios. We aim to have 2 training scenarios; each with a varying level of terrain and rain. Testing would be done on 4 other environments, with random changes to the location and levels for terrain and rain. All the models would be trained on the Jeep Wrangler model available on CARLA.

4 Metrics / Deliverables

Quantitatively, a set of performance metrics and statistical tests will be employed to evaluate and compare the results of the three agents. The following metrics will be used:

- **Success rate:** The percentage of successful waypoint following tasks completed by each agent.
- **Average completion time:** The average time it takes each agent to complete a waypoint following task.
- **Average number of collisions:** The average number of collisions per task for each agent.
- **Average deviation from the desired path:** The average deviation from the desired path for each agent.

These metrics will be used to compare the performance of each agent and determine which agent is most suitable for autonomous waypoint following in adverse conditions and varying terrain. Furthermore, statistical tests such as t-tests and ANOVA will be performed to assess the significance of the performance differences between the three agents.

The results of the quantitative analysis will provide a clear and impartial evaluation of the performance of each agent, and will be used to draw conclusions about the capabilities and limitations of each agent for autonomous waypoint following in adverse weather conditions and challenging terrain. Qualitatively, the results will be visualized by plots and videos.

Plots:

- Cross-track error plots will be compared, showcasing the total cross-track error (deviation from the desired path) and average cross-track for all three control algorithms.
- Average completion time for each agent.

- Success rate (the percentage of successful waypoint following tasks completed by each agent) of the three control algorithms.

Videos:
Animated snippets will be provided to illustrate the behavior and performance of each of the three control algorithms in the CARLA simulation environment. These visual aids will give a clear understanding of the dynamics of each agent, the challenges it encounters, and the solutions it offers.

References

Sébastien M. R. Arnold, Praateek Mahajan, Debajyoti Datta, Ian Bunner, and Konstantinos Saitas Zarkias. learn2learn: A library for meta-learning research. 8 2020.

Ignasi Clavera, Jonas Rothfuss, John Schulman, Yasuhiro Fujita, Tamim Asfour, and Pieter Abbeel. Model-based reinforcement learning via meta-policy optimization. 9 2018.

Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. Carla: An open urban driving simulator. 11 2017.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. 3 2017.

Ashish Kumar, Zipeng Fu, Deepak Pathak, and Jitendra Malik. Rma: Rapid motor adaptation for legged robots. 7 2021. URL `http://arxiv.org/abs/2107.04034`.

Óscar Pérez-Gil, Rafael Barea, Elena López-Guillén, Luis M. Bergasa, Carlos Gómez-Huélamo, Rodrigo Gutiérrez, and Alejandro Díaz-Díaz. Deep reinforcement learning based control for autonomous vehicles in carla. *Multimedia Tools and Applications*, 81:3553–3576, 1 2022. ISSN 1380-7501. doi: 10.1007/s11042-021-11437-3. URL `https://link.springer.com/10.1007/s11042-021-11437-3`.