Deep reinforcement learning for autonomous driving in AWS Deep Racer

AWS Deep Racer

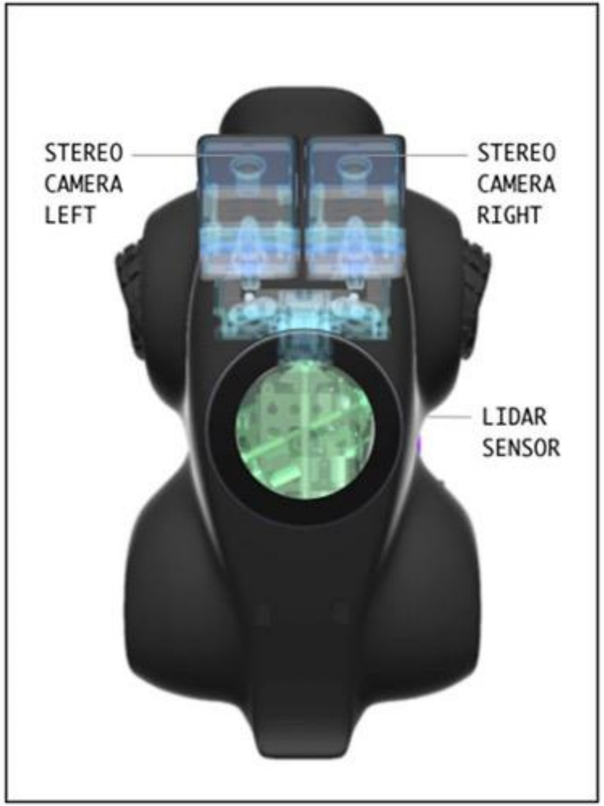
Objectives

- Time trial
- Obstacle avoidance
  - Fixed
  - Randomly placed
- Head-to-head racing

Source: https://docs.aws.amazon.com/deepracer/

Sensors

- Single camera
- Stereo camera
- + Lidar



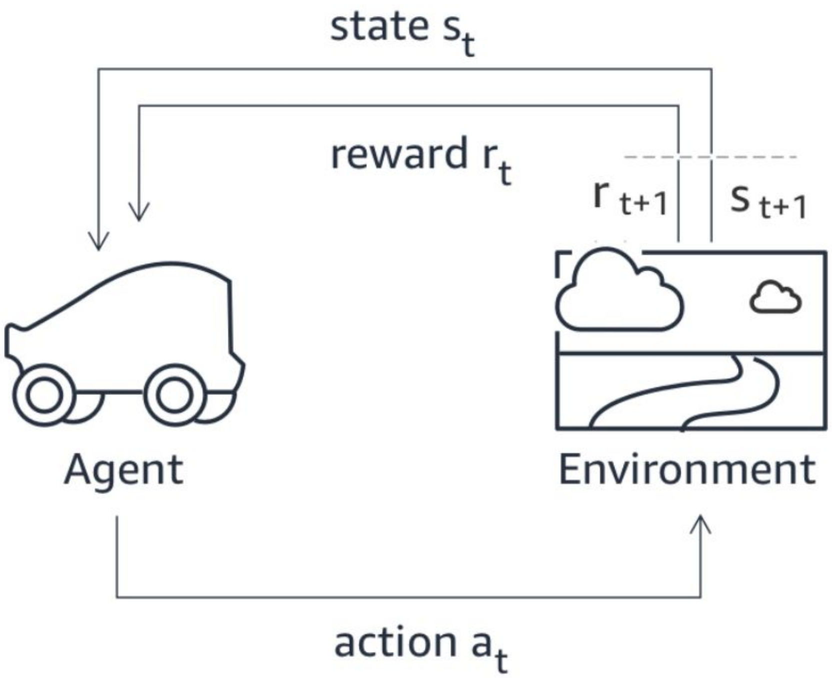
Source: https://docs.aws.amazon.com/deepracer/

Optimization algorithms

- Proxima Policy Optimization (PPO)
  - Discrete/Continuous action spaces
- Soft Actor-Critic (SAC)
  - Only continuous action space
  - Must better overcome local minima

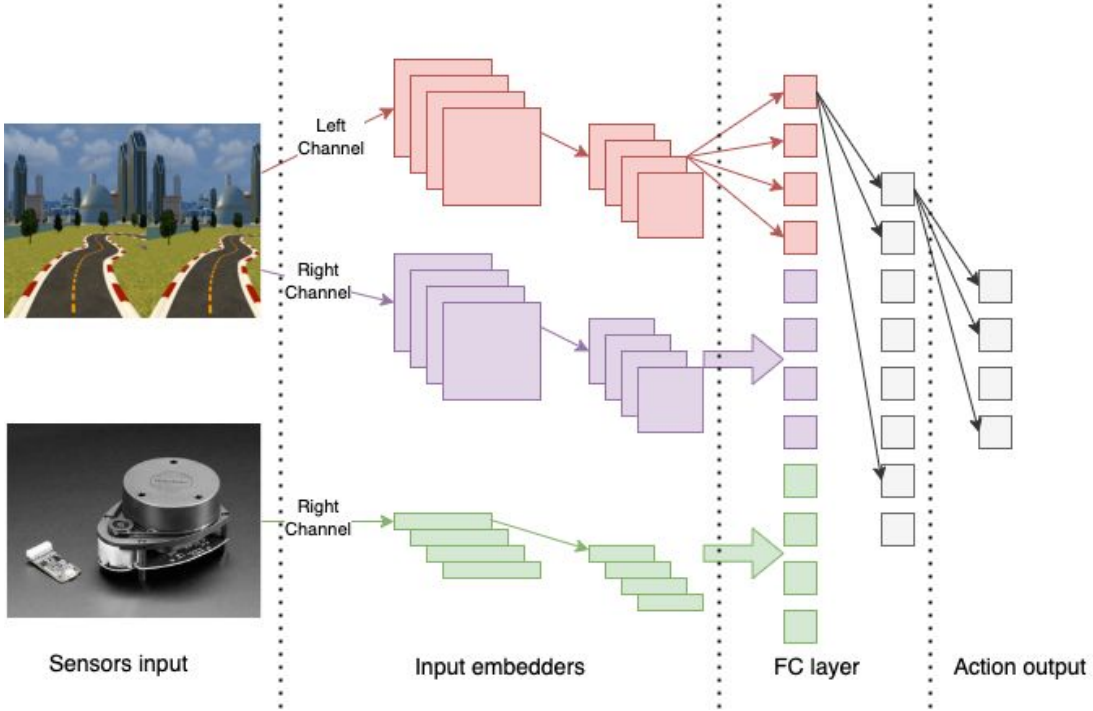
Reward function

- Takes environment parameters as input
    - Car position on the road
    - Distance from the center
    - Distance to objects
    - ...
- Returns reward (0-∞)



Source: https://docs.aws.amazon.com/deepracer/

Policy network

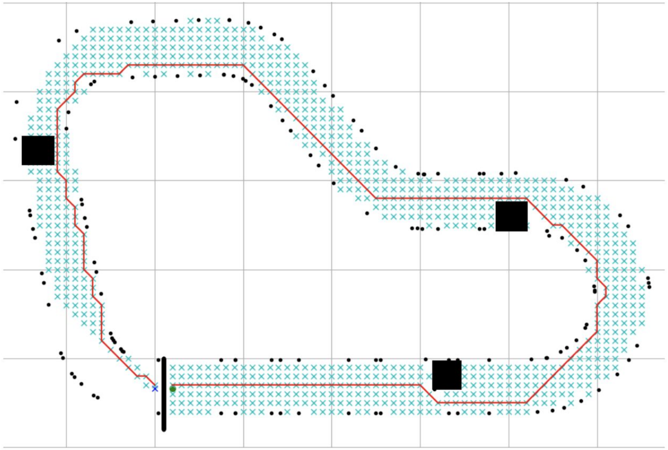


Original image

Existing solutions

Existing solutions

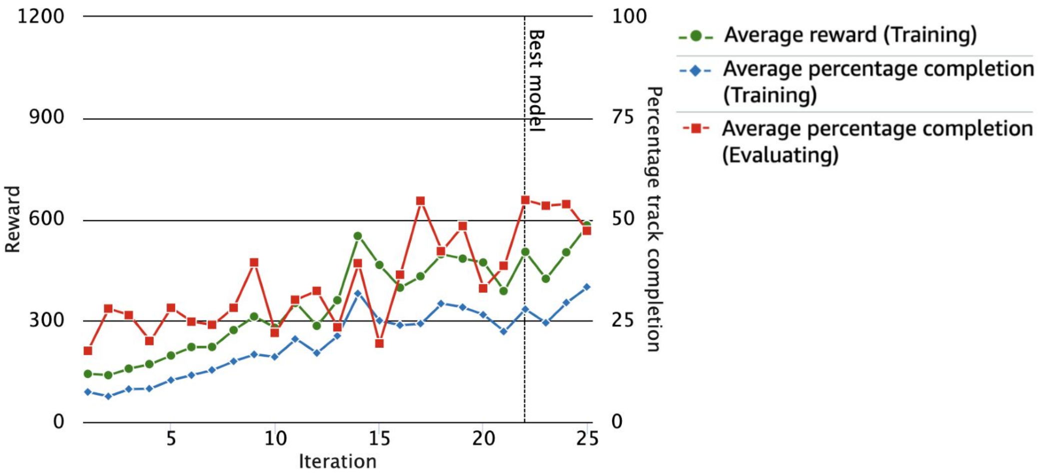
- There are not many scientific publications on this topic
- Most of them are focusing on Time trial or Fixed obstacles avoidance
- Many of the solutions don’t leverage the full power of RL



Experiments

Baseline model

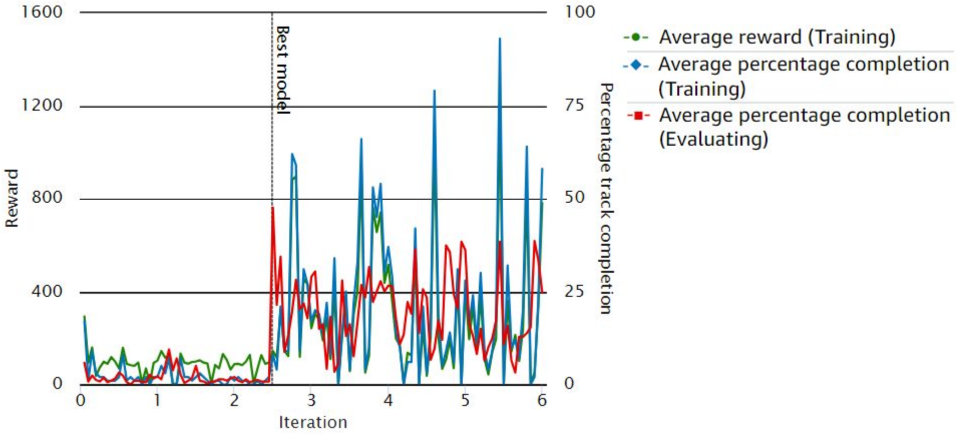
- Example reward function
- PPO
- Stereo camera



Trial	Time (MM:SS.mmm)	Trial results	Off-track	Off-track penalty	Crashes	Crash penalty
1	03:17.583	100%	0	--	26	130 seconds
2	02:09.535	100%	0	--	16	80 seconds
3	01:04.011	100%	0	--	6	30 seconds

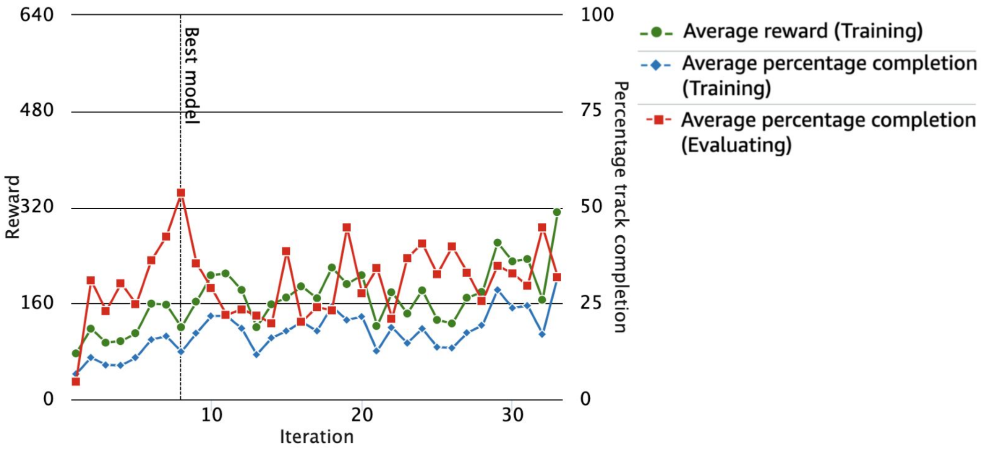
Baseline model with SAC

- Example reward function
- SAC
- Stereo camera



Reduced action space

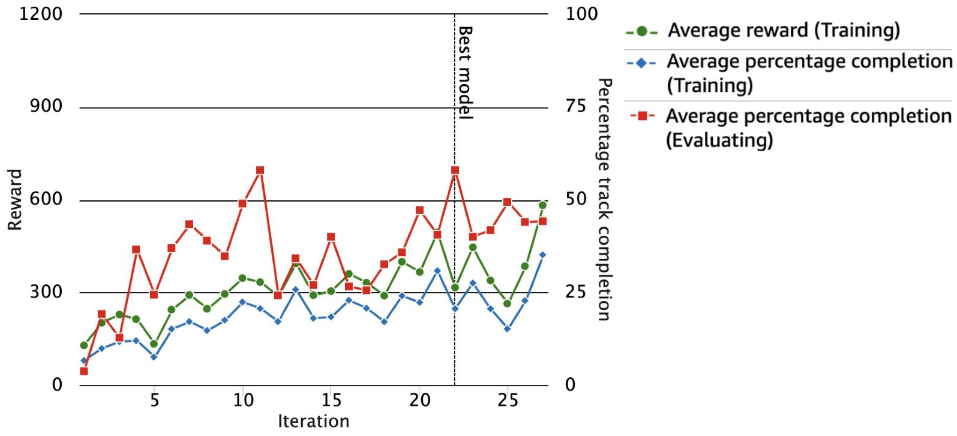
- Example reward function
- PPO
- Stereo camera
- Reduced action space



No.	Steering angle (°)	Speed (m/s)
1	-30.0	0.75
2	-15.0	0.75
3	0.0	0.75
4	15.0	0.75
5	30.0	0.75

Extended baseline model

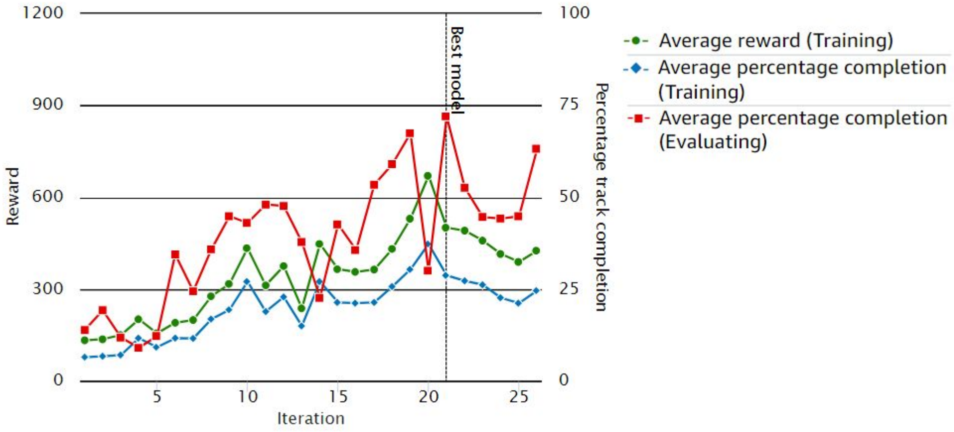
- Extended reward function
  - Discourage agent from approaching obstacles from the side
- PPO
- Stereo camera



Trial	Time (MM:SS.mmm)	Trial results	Off-track	Off-track penalty	Crashes	Crash penalty
1	00:46.98	100%	0	--	4	20 seconds
2	01:02.74	100%	1	2 seconds	6	30 seconds
3	00:52.74	100%	0	--	5	25 seconds

Extended baseline model with Lidar

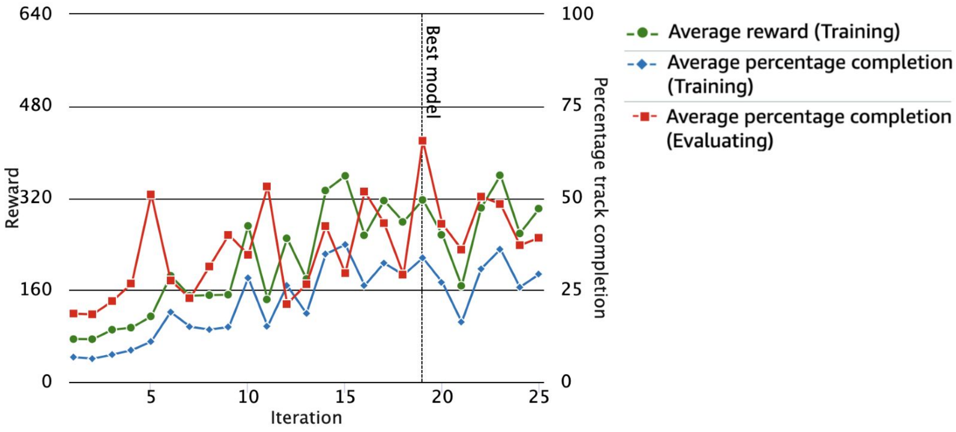
- Extended reward function
- PPO
- Singe camera + Lidar



Trial	Time (MM:SS.mmm)	Trial results	Off-track	Off-track penalty	Crashes	Crash penalty
1	00:51.988	100%	0	--	4	20 seconds
2	00:29.824	100%	0	--	1	5 seconds
3	01:04.653	100%	1	2 seconds	5	25 seconds

Continuous reward function

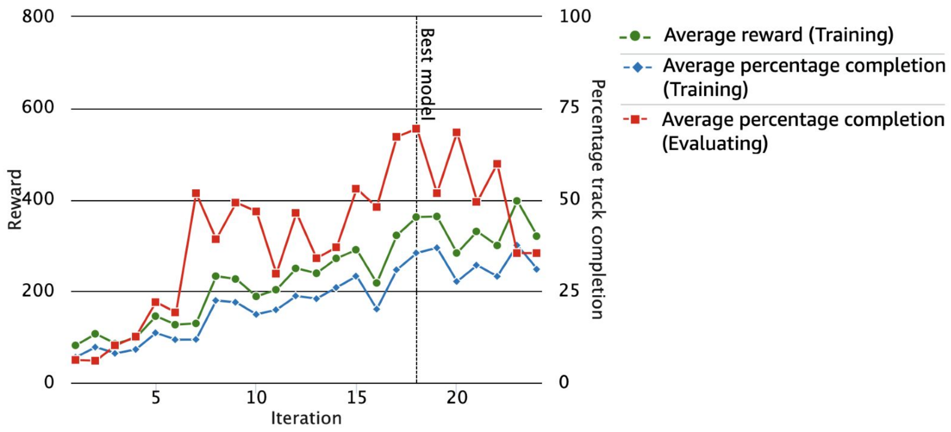
- Continuous reward function
  - Returns continuous rewards
  - Differentiable
- PPO
- Stereo camera



Trial	Time (MM:SS.mmm)	Trial results	Off-track	Off-track penalty	Crashes	Crash penalty
1	01:12.190	100%	0	--	8	40 seconds
2	00:46.606	100%	0	--	4	20 seconds
3	00:47.126	100%	1	2 seconds	4	20 seconds

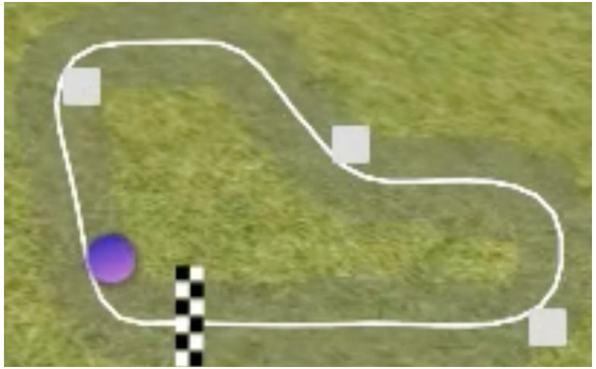
Continuous reward function with Lidar

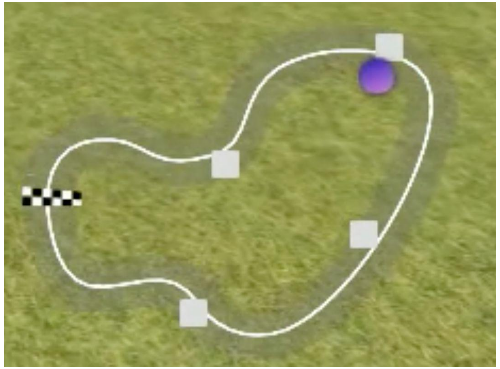
- Continuous reward function
- PPO
- Singe camera + Lidar



Trial	Time (MM:SS.mmm)	Trial results	Off-track	Off-track penalty	Crashes	Crash penalty
1	00:22.857	100%	0	--	0	--
2	00:29.673	100%	0	--	1	5 seconds
3	00:30.285	100%	1	--	1	5 seconds

Unknown environment





Trial	Time (MM:SS.mmm)	Trial results	Off-track	Off-track penalty	Crashes	Crash penalty
1	00:53.292	100%	1	2 seconds	3	15 seconds
2	00:27.494	100%	2	4 seconds	0	--
3	00:39.144	100%	2	4 seconds	0	--

Comparison

Experiment Name	Total Crashes	Total Off-track	Total Time	Laps without Crashes
Baseline	48	0	06:31.129	0
Extended Baseline	15	1	02:40.256	0
Extended Baseline with LIDAR	10	1	02:26.465	0
Continuous Reward	16	1	02:45.922	0
Continuous Reward with LIDAR	2 (⬇️95.8%)	0	01:22.815 (⬇️78.9%)	1
Continuous Reward with LIDAR, unknown environment	3	5	01:59.930	2

Publication

information

MDPI

Article

Deep Reinforcement Learning for Autonomous Driving in AWS DeepRacer

Bohdan Petryshyn [1], Serhii Postupaiev [1], Soufiane Ben Bari [1] and Armantas Ostreika [2]

[1] Department of Multimedia Engineering, Faculty of Informatics, Kaunas University of Technology, 51368 Kaunas, Lithuania
[2] Affiliation 2; e-mail@e-mail.com
* Correspondence: e-mail@e-mail.com; Tel.: (optional; include country code; if there are multiple corresponding authors, add author initials)

Abstract: The development of autonomous driving models through reinforcement learning has gained significant traction. However, developing obstacle avoidance systems remains a challenge. Specifically, optimizing path completion times while navigating obstacles is an underexplored research area. Amazon Web Services (AWS) DeepRacer emerges as a powerful infrastructure for engineering and analyzing autonomous models, providing a robust foundation for addressing these complexities. This research investigates the feasibility of training end-to-end self-driving models focused on object avoidance using reinforcement learning on the AWS DeepRacer autonomous race car platform. A comprehensive literature review of autonomous driving methodologies and machine learning model architectures is conducted, with a particular focus on object avoidance, followed by hands-on experimentation and analysis of training data. Furthermore, the impact of sensor choice, reward function, action spaces, and training time on the autonomous obstacle avoidance task are compared. Experimentation results of the best configuration demonstrate a significant improvement in obstacle avoidance performance when compared to the baseline configuration, with

IO

Information Editorial Office <information@mdpi.com>
To: Bohdan Petryshyn
Cc: Information Editorial Office <information@mdpi.com>

Dear Mr. Petryshyn,

We are writing to let you know that we have received the below submission to Information for which you are listed as a co-author.

Manuscript ID: information-2841276
Type of manuscript: Article
Title: Deep Reinforcement Learning for Autonomous Driving in AWS DeepRacer
Authors: Bohdan Petryshyn, Serhii Postupaiev, Soufiane Ben Bari, Armantas Ostreika *
Received: 9 Jan 2024

In order to confirm your connection to this submission, please click here to confirm your co-authorship:
https://susy.mdpi.com/author/confirm/931044/KZ06cW4D

Kind regards,
Information Editorial Office

Questions?