LearnyMcLearnface – Proposal

## Group Members

* Albert Anguera Sempere (albert.anguerasempere@uzh.ch)
* Dominik Bucher (dominik.bucher@uzh.ch)
* Michael Ziörjen (michael.zioerjen@uzh.ch)

This is the proposal for the project of the group **LearnyMcLearnface** for the Reinforcement Learning course in the fall semester 2020.

1. Problem

We are going to tackle the 'CarRacing-v0'[[1]](#footnote-1) problem of OpenAI by Oleg Klimov. It is a simple continuous control problem. The environment is already provided[[2]](#footnote-2):

* States consisting of 96x96 pixels with RGB dimensions (discrete 0 to 255)
* Actions consisting of “steer”, “accelerate”, “brake” (continuous -1 to 1 respectively 0 to 1)
* Rewards -0.1 every frame and +1000/N for every tile visited with N being total number of tiles in track
* Game is solved when the agent gets 900+ points consistently
* Episode finishes when agent visited all tiles or leaves boundaries

# Proposed Solutions

## (Deep Deterministic) Policy Gradient Method

We are going to use this method as a start to get familiar with the problem. As suggested during our initial discussion with Prof. Earle we are going to discretize the action space and reduce it to fewer choices (e.g., 10 steering, braking and accelerating choices). Furthermore, we can reduce the complexity of the state by turning it from RGB to grayscale using scikit-image[[3]](#footnote-3).

## 2.1 Deep Q Learning

We are going to use a neural network to approximate the Q table. As discussed in our first meeting we could simplify the problem by outputting either steering or accelerating / braking and not both at the same time.

1. https://gym.openai.com/envs/CarRacing-v0/ [↑](#footnote-ref-1)
2. https://github.com/openai/gym/blob/master/gym/envs/box2d/car\_racing.py [↑](#footnote-ref-2)
3. https://scikit-image.org/docs/stable/auto\_examples/color\_exposure/plot\_adapt\_rgb.html [↑](#footnote-ref-3)