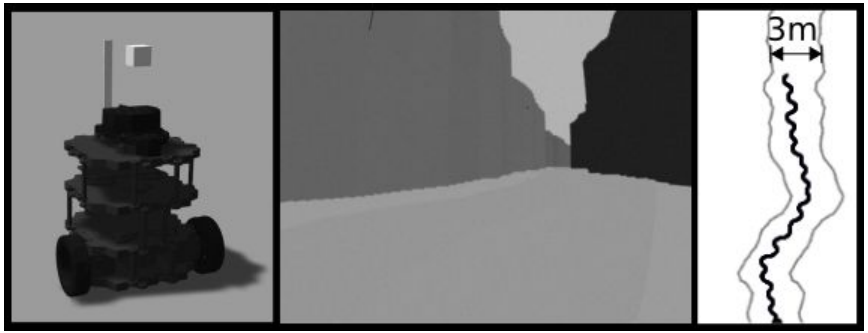
Future Depth as Value Signal for Learning Collision Avoidance

Klaas Kelchtermans, Tinne Tuytelaars

KU LEUVEN

Problem Statement:



Monocular collision avoidance

A good RL algorithm:

+ Self-supervised reward
+ Off-Policy
+ Simple

Inspired on the work of *Self-supervised Deep Reinforcement Learning with Generalized Computation Graphs for Robot Navigation, G. Kahn et al. ICRA2018*

Improvement:

+ No actual collisions
+ Better performance
+ Less data required

Method

$$V_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots$$
$$V_t = D_{t+1} - D_t + \gamma(D_{t+2} - D_{t+1}) + \gamma^2(D_{t+3} - D_{t+2}) + \dots$$
$$V_t = -D_t + (1-\gamma)D_{t+1} + \gamma(1-\gamma)D_{t+2} + \dots$$
$$V_t \approx -D_t$$

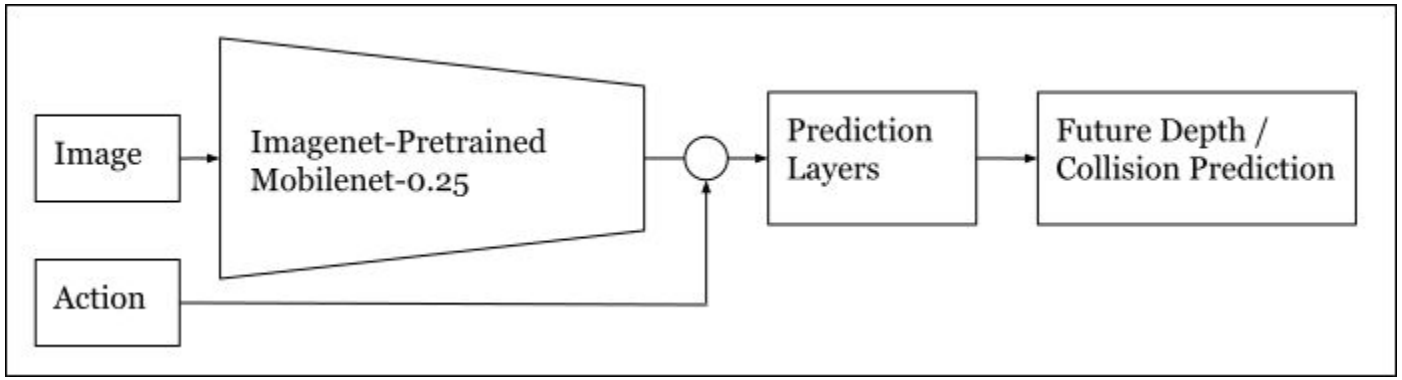
$$D_t^{'} = r_t + D_{t+1}^{'} \Rightarrow r_t = D_t^{'} - D_{t+1}^{'}$$

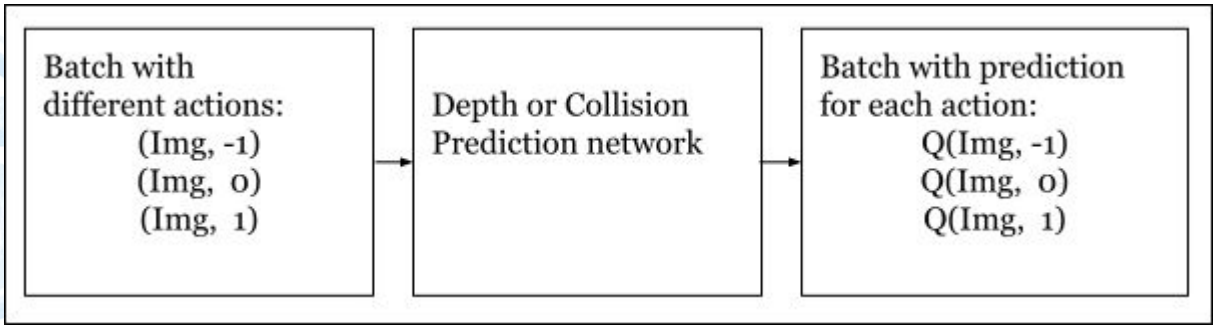
$$Q(I_t, a_t) = r_t = D_{t+1}$$

$$\pi(I_t) = argmax_{a_t}(min(D_{t+1}(I_t, a_t)))$$

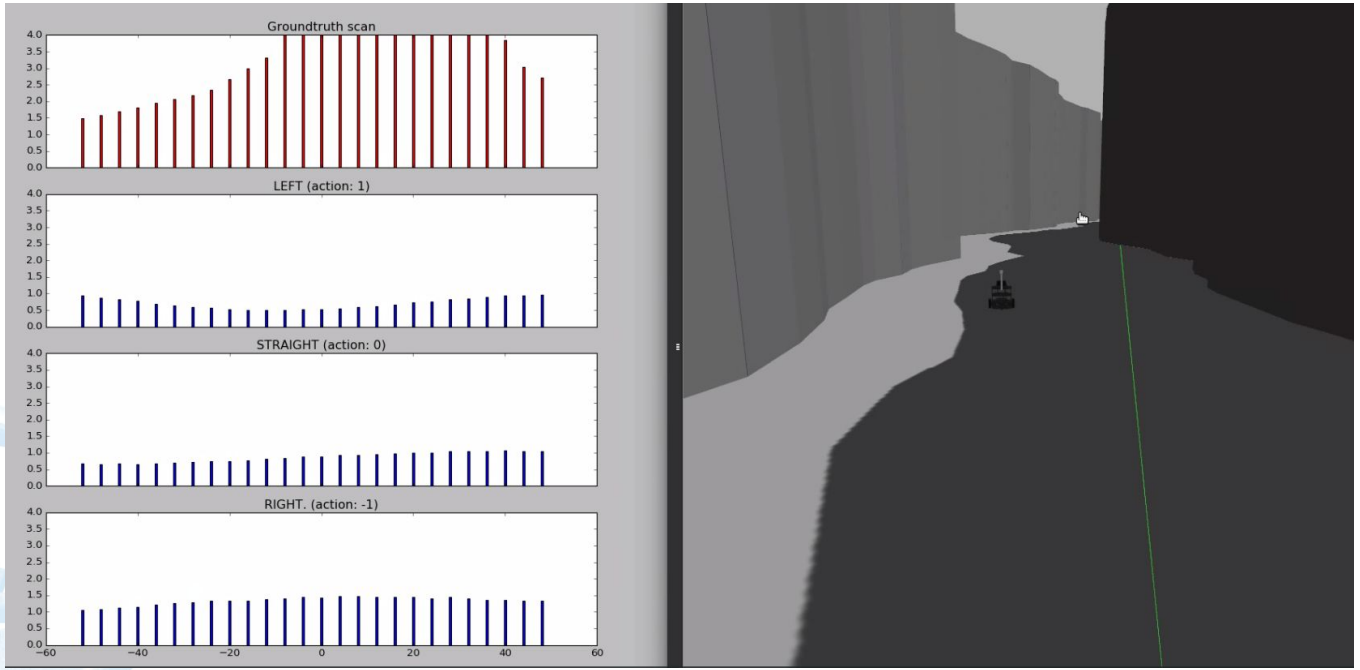
KU LEUVEN

Implementation



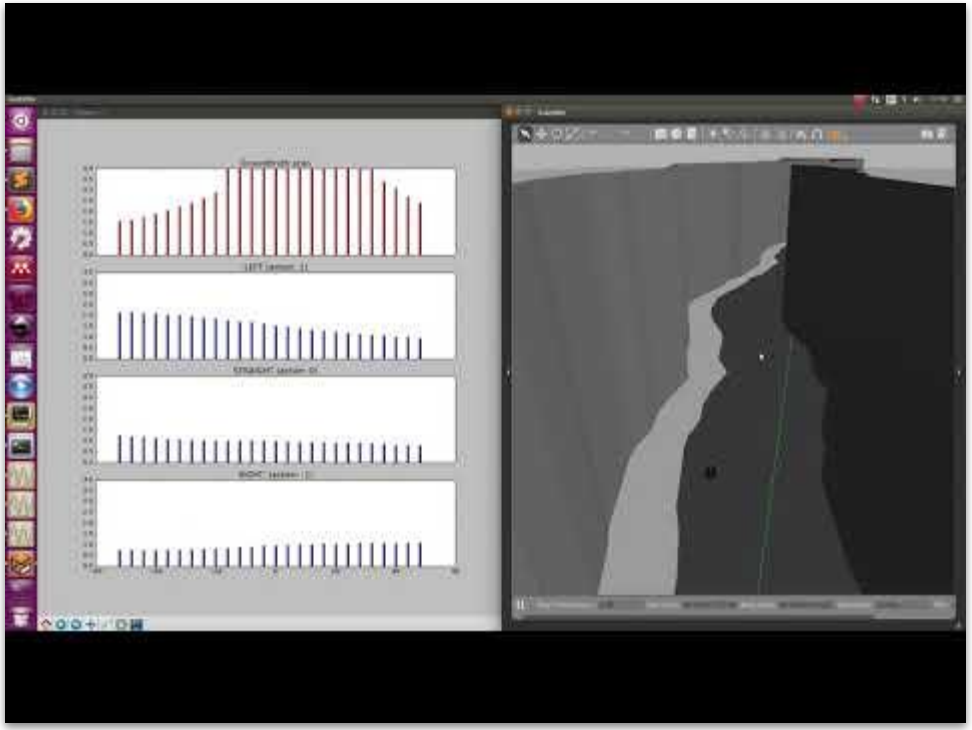


Results



KU LEUVEN

Results



KU LEUVEN