

Reinforcement Learning Project

Rubber Duck Racing

Jim Rhotert



The CarRacing-v0 Environment

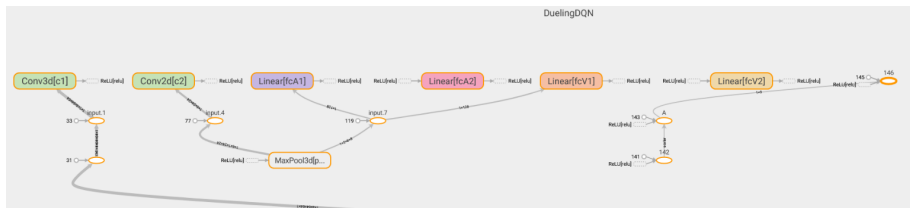
- Learning from 96x96 pixels
- Reward function:
 - ▶ -0.1 every frame
 - ▶ $+1000 / N$ for every track tile visited
 - ▶ Driving into the void: -100 penalty



- Framestacking 4 frames to capture movement
- Greyscaling
- Resizing observations to 64x64
- Discrete action space 1x9:
 - ▶ Gas (50%)
 - ▶ Gas + Direction
 - ▶ Brake
 - ▶ Brake + Direction
 - ▶ Pure Turn (both directions)
 - ▶ Nothing

Our Agent

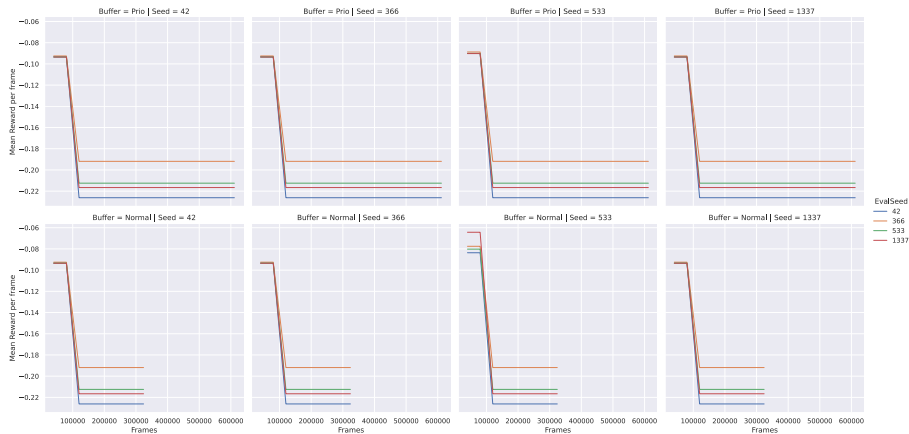
- DoubleDuelingDQN
- Polyak Averaging with $\tau = 0.01$
- Prioritized Replay Buffer / Replay Buffer
- ε -greedy exploration - linear decay from 1 to 0.1 over 1M frames
- Learning rate $\alpha = 0.01$
- Discount Factor $\gamma = 0.9$



Training Visualization



Evaluation of Training Checkpoints



Thanks for listening!