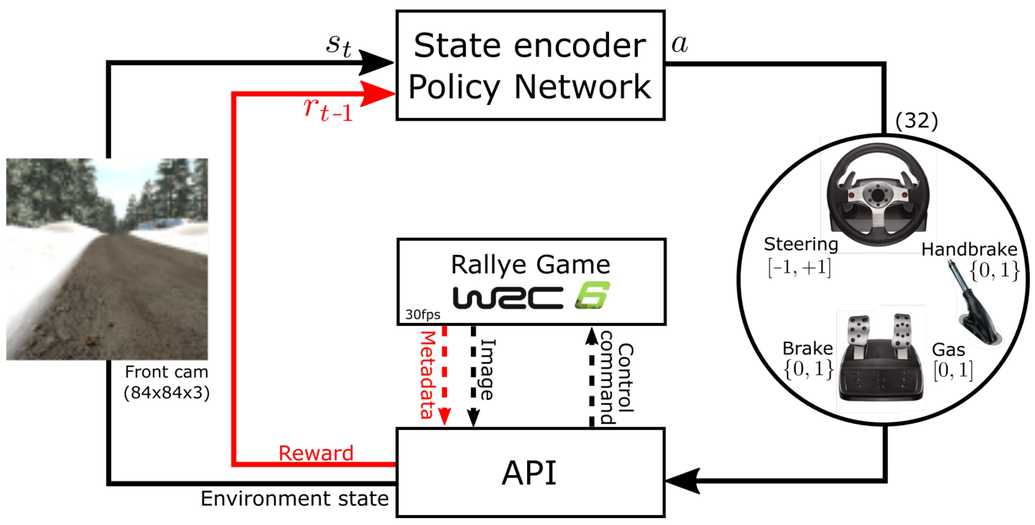
End-to-End Race Driving with Deep Reinforcement Learning

[1] Inria informatics mathematics — [2] Valeo

Maximilian Jaritz[1,2] Raoul de Charette[1] Marin Toromanoff[2] Etienne Perot[2] Fawzi Nashashibi[1]

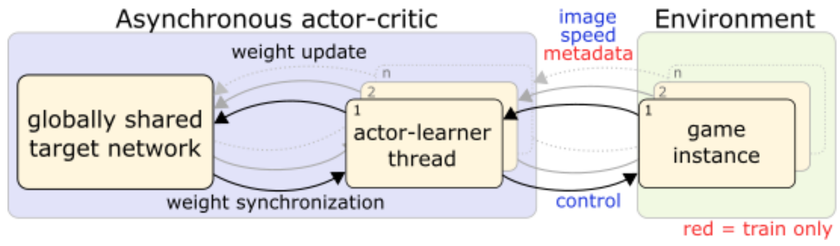


Learning direct control from RGB images

Need to learn world representation and car dynamics from scratch
Challenges: full control, realistic graphics and dynamics

Self-supervised learning

Using Asynchronous Reinforcement Learning (A3C) framework



Bots evolve in separate environments asynchronously
Achieves experience decorrelation without experience replay
Learn to maximize discounted reward from sparse reward

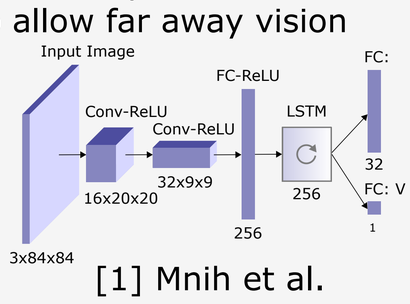
Main related works

[1] Mnih et al., Asynchronous reinforcement learning (A3C)
[2] Chen et al., Learning high level feature for *direct perception*
[3] Bojarski et al., Imitation learning from expert driver

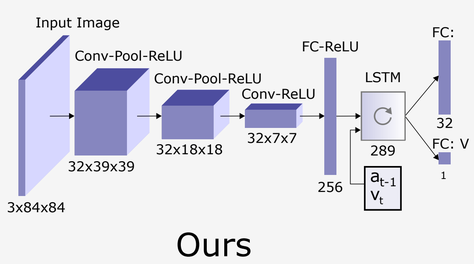
CONTRIBUTIONS

New state encoder

Use a deeper encoder with past actions and speed to allow far away vision



[1] Mnih et al.



Ours

Control strategy

Learn full control (steering, gas, brake, handbrake)
Stochastic choice of discrete commands (as good as continuous)

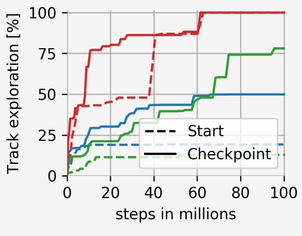
# classes	Control commands			
	Steering	Gas	Brake	Hand brake
27	$\{-1., -0.75, \ldots, 1.\}$	$\{0.0, 0.5, 1.0\}$	$\{0\}$	$\{0\}$
4	$\{-1., -0.5, 0.5, 1.\}$	$\{0.0\}$	$\{0\}$	$\{1\}$
1	$\{0.0\}$	$\{0.0\}$	$\{1\}$	$\{0\}$

Note prominence of gas commands

Respawn strategy

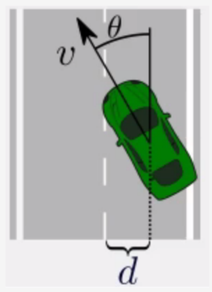
Maximize environment variance for better asynchronous learning

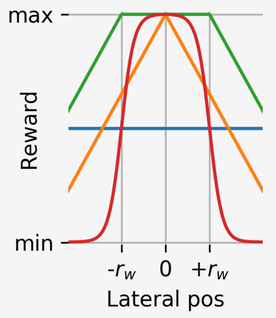
[1] Mnih et al. Ours

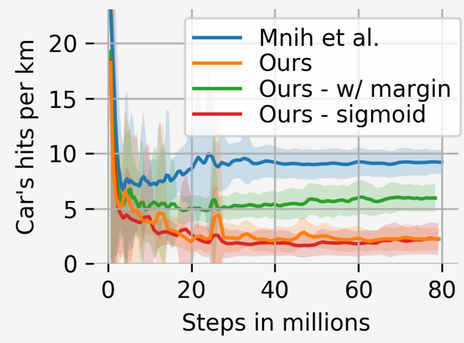


Reward shaping

New frame-wise rewards
Enforce smooth reward using road distance as a penalty

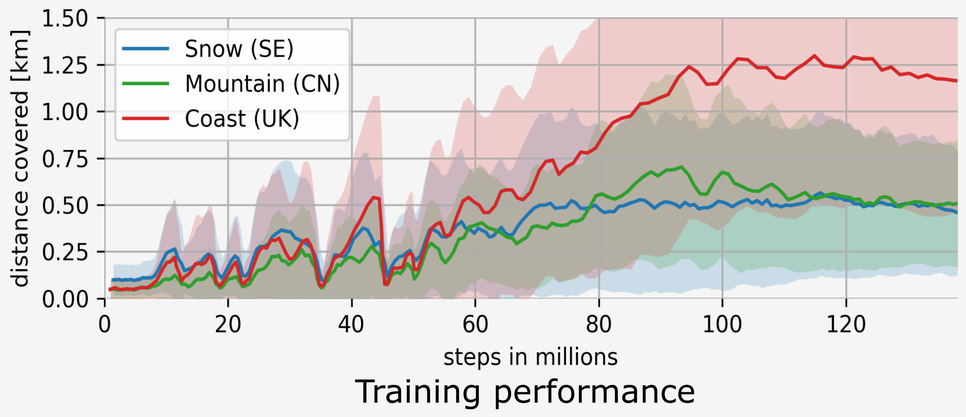






Training setup

Trained asynchronously with 15 agents (on 3 PC)
Simultaneously driving on 3 tracks (mountains, snow, coast)

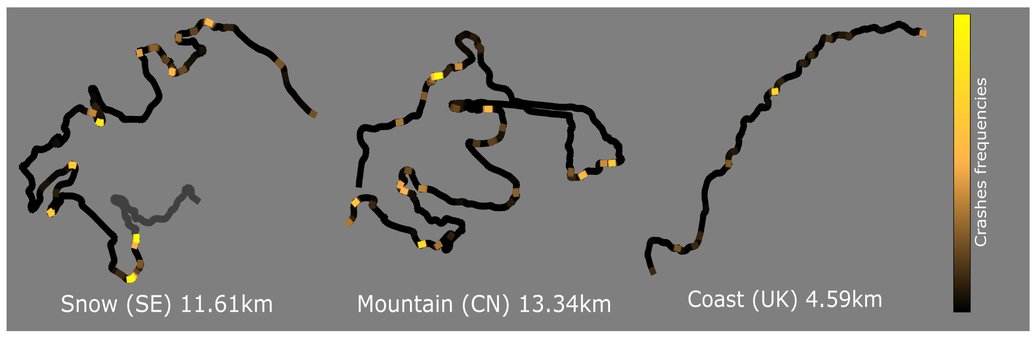


Training performance

Gaming performance

Study of training tracks driving style

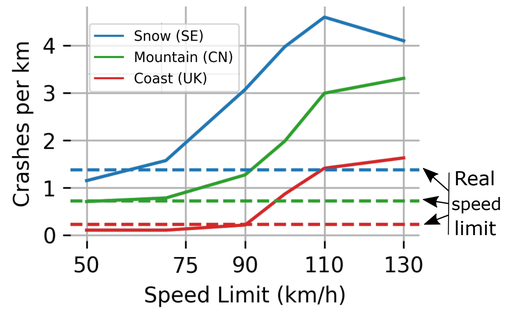
Learned to drive at 72.9km/h and covers 0.72km per run
Driving style is significantly smoother



Training generalized to unseed tracks (check the video)

Influence of speed limits

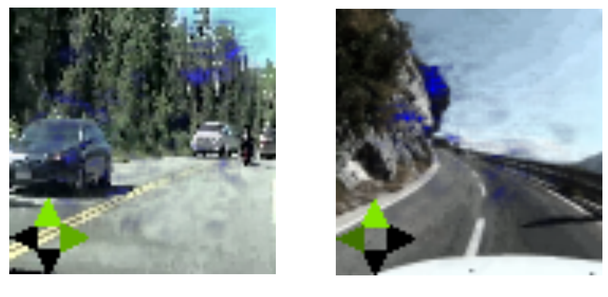
Use road curvature to limit speed from real road design (dashed lines)



Real data performance

Prediction on real videos

Tested on real videos (web-footage) using open-loop predictions



Check out the videos

https://team.inria.fr/rits/drl/

[1] Mnih et al., Asynchronous methods for deep reinforcement learning, ICML 2016
[2] Chen et al., DeepDriving: Learning affordance for direct perception in autonomous driving, ICCV 2015
[3] Bojarski et al., End to end learning for self-driving cars, arxiv 2016