GRIAD: General Reinforced Imitation for Autonomous Driving

R. Chekroun, M. Toromanoff, S. Hornauer, F. Moutarde

MINES ParisTech - PSL University & Valeo Driving Assistance Research

Introduction

Introduction

$\rightarrow$ Based on a novel method: General Reinforced Imitation (GRI) by Chekroun et al. [1]

Introduction

$\rightarrow$ Based on a novel method: General Reinforced Imitation (GRI) by Chekroun et al. [1]

$\rightarrow$ GRI-based algorithm in the CARLA Challenge 2021:

- Ranked #1 in the MAP track.
- Ranked #4 in the SENSOR track.

Introduction

→ Based on a novel method: General Reinforced Imitation (GRI) by Chekroun et al. [1]

→ GRI-based algorithm in the CARLA Challenge 2021:

- Ranked #1 in the MAP track.
- Ranked #4 in the SENSOR track.

→ Focusing GRIAD for camera-based end-to-end autonomous driving.

General Reinforced Imitation

The GRI method

$\rightarrow$ Combines benefits from exploration and expert data.

The GRI method

$\rightarrow$ Combines benefits from exploration and expert data.

$\rightarrow$ Simplifying hypothesis: expert demonstrations can be seen as perfect data whose underlying policy gets a constant high reward.

The GRI method

$\rightarrow$ Combines benefits from exploration and expert data.

$\rightarrow$ Simplifying hypothesis: expert demonstrations can be seen as perfect data whose underlying policy gets a constant high reward.

$\rightarrow$ Defines two types of agents:

- The exploration agent: gather data by exploring the environment.
- The demonstration agent: send pre-generated expert data associated with a demonstration reward.

GRI: the algorithm

Algorithm 1: GRI: General Reinforced Imitation

Input: r_{demo} demonstration reward value, p_{demo} probability to use demonstration agent;
Initialize empty buffer $\mathcal{B}$;
while *not converged* **do**
    if *len($\mathcal{B}$) $\geq$ min_buffer* **then**
        do a DRL network update;
    end
    if *random.random()* $\geq p_{demo}$ **then**
        collect episode $(s_t^{online}, a_t, r_t, s_{t+1}^{online})_t$ in buffer $\mathcal{B}$ with exploration agent
    else
        add episode $(s_t^{offline}, a_t, r_{demo}, s_{t+1}^{offline})_t$ in buffer $\mathcal{B}$ with demonstration agent;
    end
end

Figure 1: GRI algorithm

GRI for Autonomous Driving

A modular end-to-end pipeline (i)

→ Pipeline inspired by Toromanoff et al. IAs method [6].

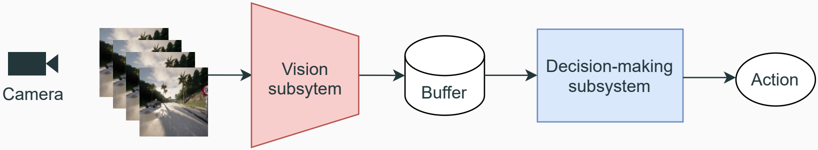
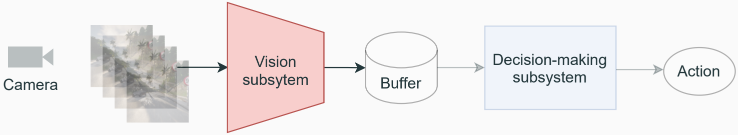


Figure 2: Modular pipeline for end-to-end autonomous driving using reinforcement learning. It is composed of a vision subsystem and a decision-making subsystem.

A modular end-to-end pipeline (ii)

Design of the vision subsystem



Design of the vision subsystem

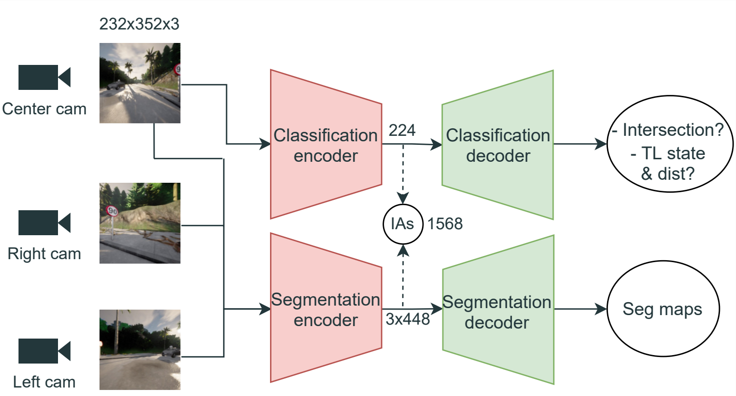
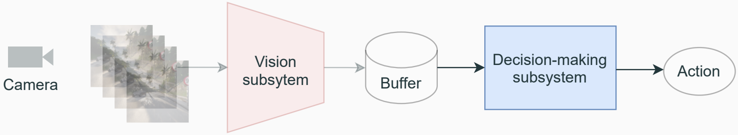


Figure 3: Two encoder-decoder networks are pretrained on segmentation, classifications and regression tasks. After training, the visual encoders serve as fixed feature extractors with frozen weights. For the DRL backbone training, both encoder outputs are concatenated and sent to the memory buffer as input to DRL.

A modular end-to-end pipeline (iii)

Design of the decision-making subsystem



Design of the decision-making subsystem

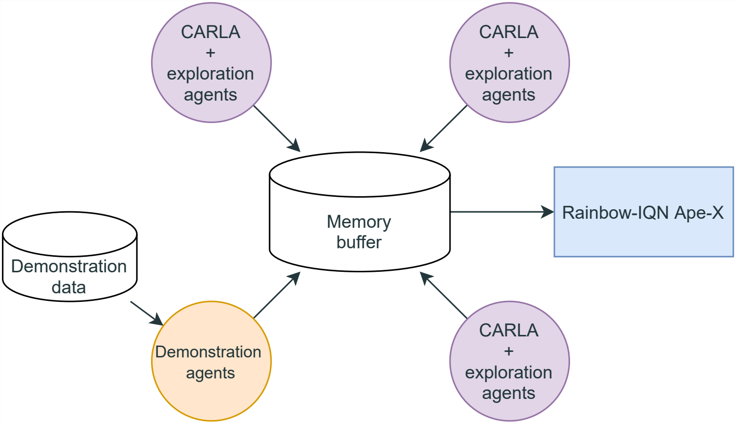


Figure 4: Simplified representation of the distributed GRIAD setup with a Rainbow-IQN Ape-X [5] backbone. A central computer receives data in a shared replay buffer from both exploration and demonstration agents running on other computers.

Representation of the whole system

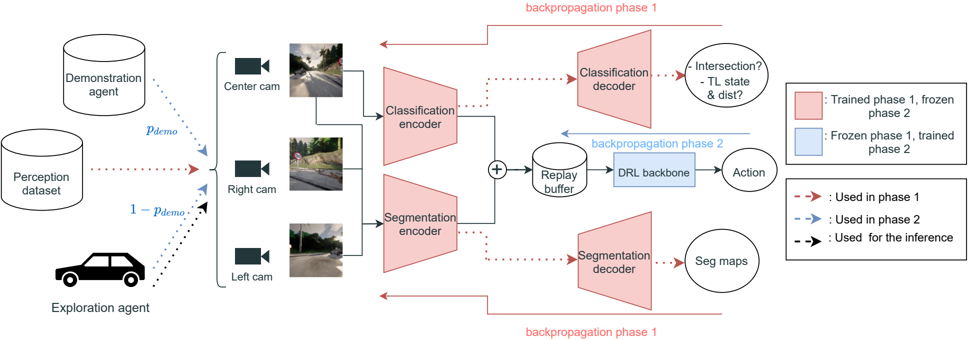


Figure 5: This pipeline is trained in two phases: (1) Visual encoders are pretrained on a perception dataset on several auxiliary tasks. (2) Visual encoders are frozen and a GRI-based DRL network is trained with both pre-generated expert data with an offline demonstration agent and an online exploration agent gathering data from a simulator.

Experimental results

Ablation study on the NoCrash benchmark

Task	Town, Weather	RL 12M	RL 16M	GRIAD
Empty	train, train	96.3 ± 1.5	**98.0 ± 1.0**	**98.0 ± 1.7**
Regular		95.0 ± 2.4	**98.6 ± 1.2**	**98.3 ± 1.7**
Dense		91.7 ± 2.0	**95.0 ± 1.6**	93.7 ± 1.7
Empty	test, train	83.3 ± 3.7	**96.3 ± 1.7**	94.0 ± 1.6
Regular		82.6 ± 3.7	**96.3 ± 2.5**	93.0 ± 0.8
Dense		61.6 ± 2.0	**78.0 ± 2.8**	**77.7 ± 4.5**
Empty	train, test	67.3 ± 1.9	73.3 ± 2.5	**83.3 ± 2.5**
Regular		76.7 ± 2.5	81.3 ± 2.5	**86.7 ± 2.5**
Dense		67.3 ± 2.5	80.0 ± 1.6	**82.6 ± 0.9**
Empty	test, test	60.6 ± 2.5	62.0 ± 1.6	**68.7 ± 0.9**
Regular		59.3 ± 2.5	56.7 ± 3.4	**63.3 ± 2.5**
Dense		40.0 ± 1.6	46.0 ± 3.3	**52.0 ± 4.3**

Table 1: Ablation study of GRIAD using the NoCrash benchmark. GRIAD experimentally shows to generalize more on test weather than RL with 12M and 16M steps and globally gives the best agent.Mean and standard deviation over 3 evaluation seeds.

On the CARLA Leaderboard

	Cam.	DS	RC	IS
LBC [3]	3	10.9	21.3	0.55
IAs [6]	1	24.98	46.97	0.52
Rails [2]	4	31.37	57.65	0.56
GRIAD	3	**36.79**	**61.85**	**0.60**

Table 2: CARLA Leaderboard's driving metrics on camera-based systems. GRIAD was trained on CARLA 0.9.10.

On the CARLA Leaderboard

	Cam.	DS	RC	IS
LBC [3]	3	10.9	21.3	0.55
IAs [6]	1	24.98	46.97	0.52
Rails [2]	4	31.37	57.65	0.56
GRIAD	3	**36.79**	**61.85**	**0.60**

Table 2: CARLA Leaderboard's driving metrics on camera-based systems. GRIAD was trained on CARLA 0.9.10.

	Cam.	LiDAR	IMU	DS	RC	IS
TF+ [4]	1	✓	✓	37.49	74.83	0.54
TF Ens. [4]	3	✓	✓	37.84	72.36	**0.60**
LAV	4	✓	✓	**47.65**	**87.18**	0.53
GRIAD	3	✗	✗	36.79	61.85	**0.60**

Table 3: CARLA Challenge 2021 final driving metrics on the SENSOR track. GRIAD was trained on CARLA 0.9.10.

Conclusion & Additional information

$\rightarrow$ GRIAD efficiently leverages expert data and exploration for end-to-end autonomous driving.

Conclusion & Additional information

→ GRIAD efficiently leverages expert data and exploration for end-to-end autonomous driving.

→ GRI method generalizability and robustness assessed in Chekroun et al. [1]:

- Robust to noisy demonstrations
- GRI further validated on the Mujoco benchmark

Conclusion & Additional information

→ GRIAD efficiently leverages expert data and exploration for end-to-end autonomous driving.

→ GRI method generalizability and robustness assessed in Chekroun et al. [1]:

- Robust to noisy demonstrations
- GRI further validated on the Mujoco benchmark

→ Accelerates the training: sending pre-generated expert data is almost free compared to interaction with the CARLA simulator.

Conclusion & Additional information

→ GRIAD efficiently leverages expert data and exploration for end-to-end autonomous driving.

→ GRI method generalizability and robustness assessed in Chekroun et al. [1]:

- Robust to noisy demonstrations
- GRI further validated on the Mujoco benchmark

→ Accelerates the training: sending pre-generated expert data is almost free compared to interaction with the CARLA simulator.

→ To go further: add LiDaR to improve the visual subsystem.

Conclusion & Additional information

Thank you for your attention!

Do not hesitate to contact me: raphael.chekroun@minesparis.psl.eu

- Raphael Chekroun, Marin Toromanoff, Sascha Hornauer, and Fabien Moutarde.
  Gri: General reinforced imitation and its application to vision-based autonomous driving, 2021.
- Dian Chen, Vladlen Koltun, and Philipp Krähenbühl.
  Learning to drive from a world on rails.
  In *ICCV*, 2021.
- Dian Chen, Brady Zhou, Vladlen Koltun, and Philipp Krähenbühl.
  Learning by cheating.
  In *Conference on Robot Learning (CoRL)*, 2019.

Aditya Prakash, Kashyap Chitta, and Andreas Geiger.
Multi-modal fusion transformer for end-to-end autonomous driving.
In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2021.

Marin Toromanoff, Emilie Wirbel, and Fabien Moutarde.
Is Deep Reinforcement Learning Really Superhuman on Atari?
In *Deep Reinforcement Learning Workshop of 39th Conference on Neural Information Processing Systems (Neurips'2019)*, Vancouver, Canada, December 2019.

Marin Toromanoff, Emilie Wirbel, and Fabien Moutarde.
End-to-end model-free reinforcement learning for urban driving using implicit affordances.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.