Agile Autonomous Driving using End-to-End Deep Imitation Learning

Presenter: Shaoheng Fang

2023.10.03

Task: Off-Road Autonomous Driving

A challenging robotics problem:

- Physically-complex, uncertain environment
- The surface is constantly evolving and highly stochastic

Requirements:

- Precise steering and throttle maneuvers
- High speed, high-frequency decisions

Task: Off-Road Autonomous Driving

A challenging robotics problem:

- Physically-complex, uncertain environment
- The surface is constantly evolving and highly stochastic

Requirements:

- Precise steering and throttle maneuvers
- High speed, high-frequency decisions

↓

An end-to-end imitation learning system

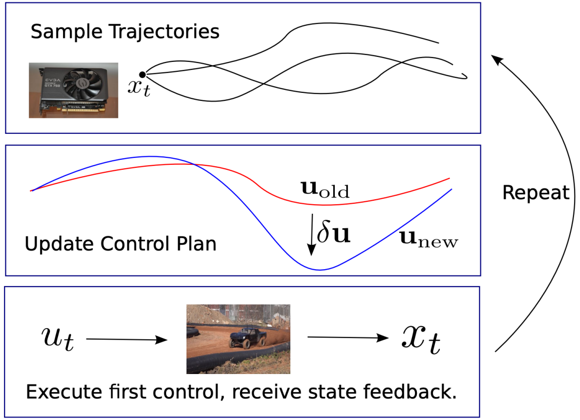
Previous Works

Model-predictive control (MPC) algorithm

An optimal control technique in which the calculated control actions minimize a cost function for a constrained dynamical system

Model Predictive Path Integral (MPPI) Control

- Sample multiple trajectories from current state
- Compute the cost function of each trajectory
- Update the control sequence
- Execute the first control, update the initial control sequence



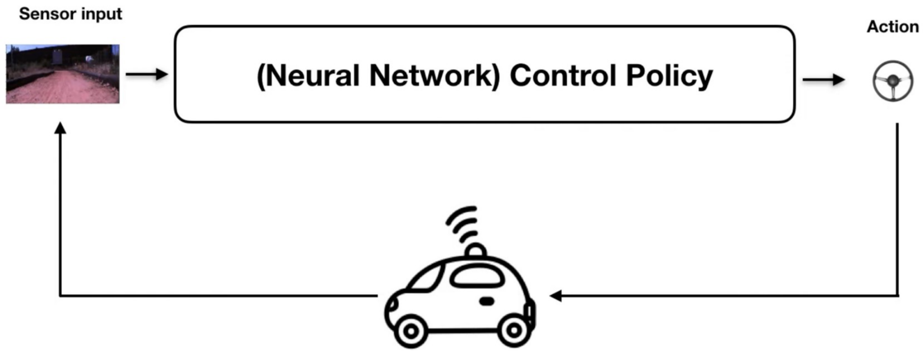
Motivation — End-to-End Learning

Previous: Model-predictive control (MPC)

- Expensive sensors (GPS, IMU) for state estimation
- High frequency online replanning (low speed)
- Computationally expensive

End-to-End framework

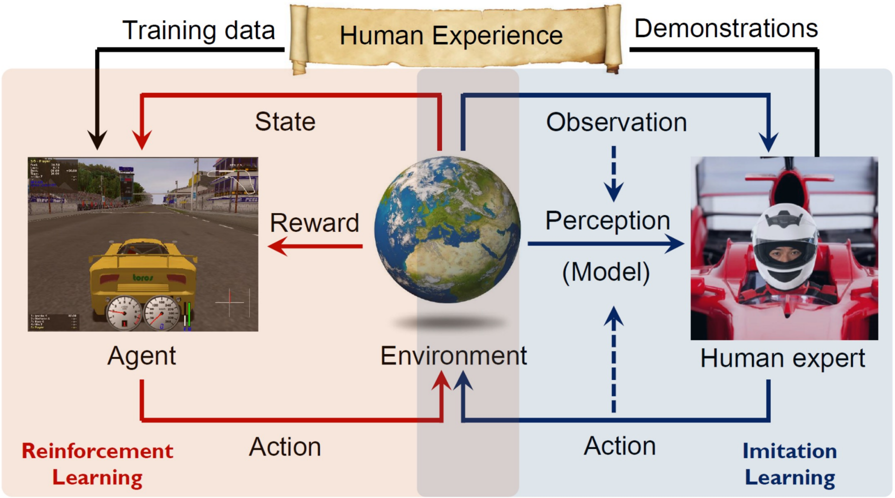
- Lower cost
- Faster inference time



Motivation — Imitation Learning

Model-free reinforcement learning:

- Sample inefficiency
- Costly
- Damage to the robots



Imitation Learning

Learn the policy from an expert

An expert is needed (human / teacher model)

Previous Works

Imitation Learning for autonomous driving

Methods	Tasks	Observations	Action	Algorithm	Expert	Experiment
[1]	On-road low-speed	Single image	Steering	Batch	Human	Real &simulated
[23]	On-road low-speed	Single image & laser	Steering	Batch	Human	Real &simulated
[24]	On-road low-speed	Single image	Steering	Batch	Human	Simulated
[20]	Off-road low-speed	Left & right images	Steering	Batch	Human	Real
[33]	On-road unknown speed	Single image	Steering + break	Online	Pre-specified policy	Simulated
Our Method	Off-road high-speed	Single image + wheel speeds	Steering + throttle	Batch & online	Model predictive controller	Real & simulated

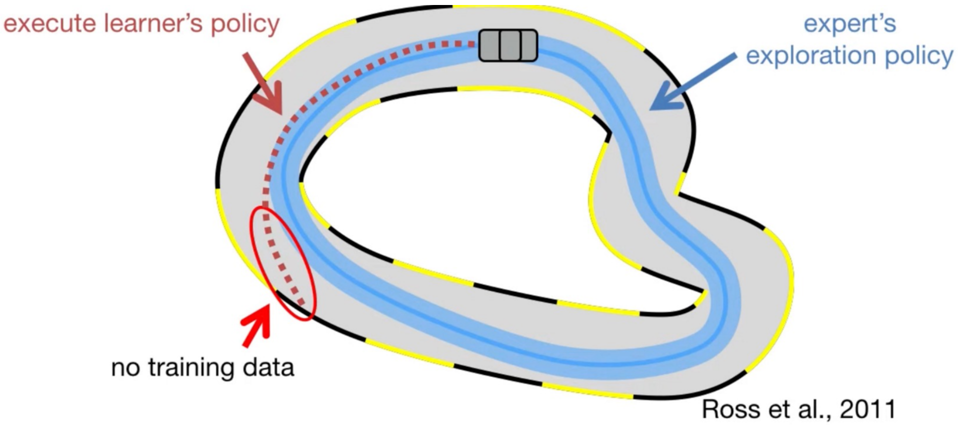
off-road; real-world experiment environment ; online imitation learning

Preliminary

Batch Imitation Learning vs Online Imitation Learning

Batch Imitation Learning:

Collect training data D from an expert policy → Train the learner's policy on D



Compound error

Preliminary

Online Imitation Learning

DAgger (Dataset Aggregation) Method

Initialize $\mathcal{D} \leftarrow \emptyset$.
Initialize $\hat{\pi}_1$ to any policy in Π.
for $i = 1$ **to** N **do**
  Let $\pi_i = \beta_i \pi^* + (1 - \beta_i)\hat{\pi}_i$.
  Sample T-step trajectories using π_i.
  Get dataset $\mathcal{D}_i = \{(s, \pi^*(s))\}$ of visited states by π_i and actions given by expert.
  Aggregate datasets: $\mathcal{D} \leftarrow \mathcal{D} \bigcup \mathcal{D}_i$.
  Train classifier $\hat{\pi}_{i+1}$ on $\mathcal{D}$.
end for
Return best $\hat{\pi}_i$ on validation.

π^*: expert policy

β: mixing rate

Experimental Setup

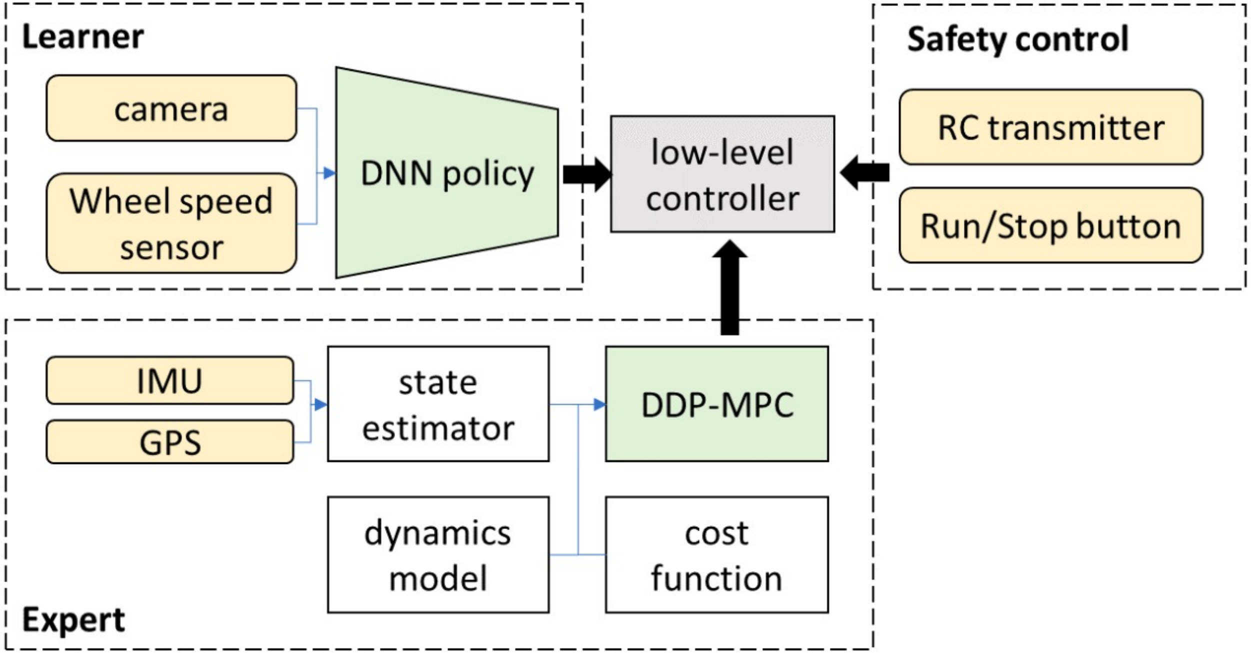
Autonomous driving platform:

- 1/5-scale autonomous AutoRally car
- On-board device (GTX 750 Ti GPU)
- Camera sensor: \$500, GPS/IMU: \$6000
- 50 Hz sampling rate

Off-road autonomous driving task

- A fixed dirt track
- A desired speed of 7.5 m/s

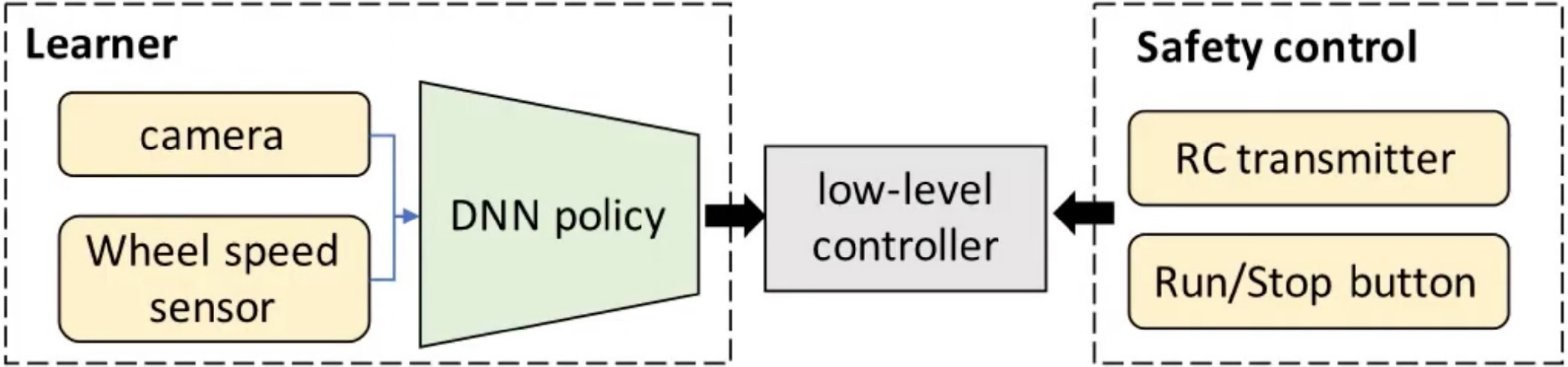
Framework



Training Phase

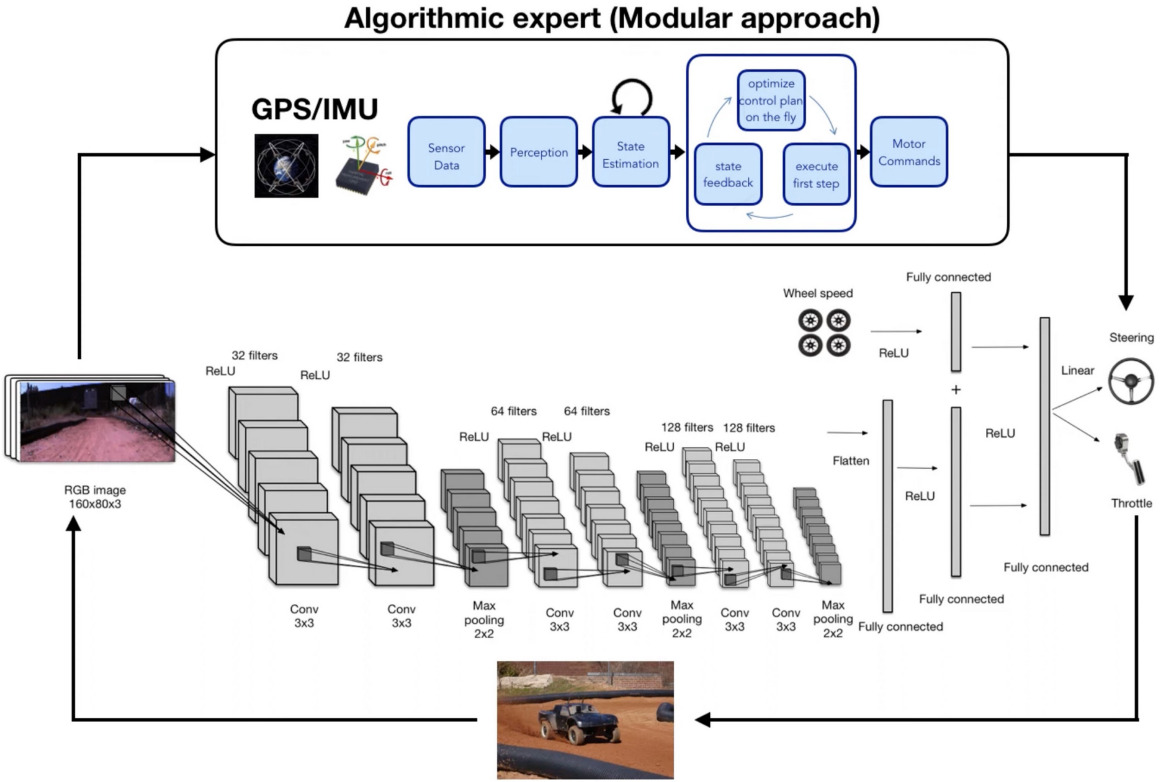
Framework

After the entire learning session of each setting, three rollouts will be performed using the learned policy for performance evaluation.



Testing Phase

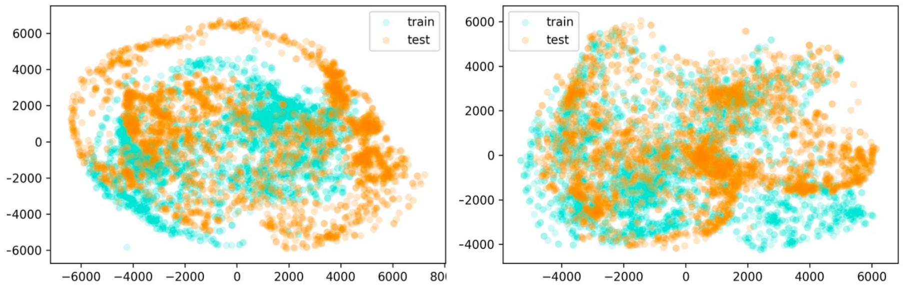
Framework



Algorithmic expert (Modular approach)
GPS/IMU
Sensor Data
Perception
State Estimation
optimize control plan on the fly
state feedback
execute first step
Motor Commands
Wheel speed
Fully connected
ReLU
Steering
Linear
Throttle
32 filters
ReLU
32 filters
ReLU
64 filters
64 filters
ReLU
ReLU
128 filters
128 filters
ReLU
ReLU
Flatten
RGB image 160x80x3
Conv 3x3
Conv 3x3
Max pooling 2x2
Conv 3x3
Conv 3x3
Max pooling 2x2
Conv 3x3
Conv 3x3
Max pooling 2x2
Fully connected
Fully connected
Fully connected

Experimental Results — Qualitative

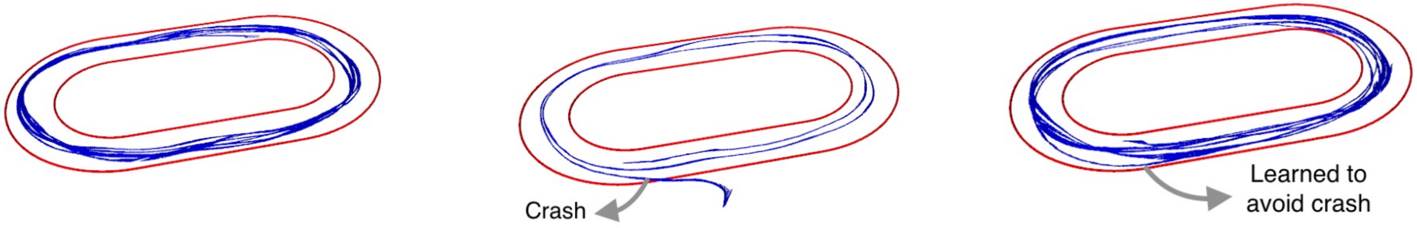
Batch IL v.s. Online IL



(a) Batch raw image (b) Online raw image

Data distritbutions (t-SNE)

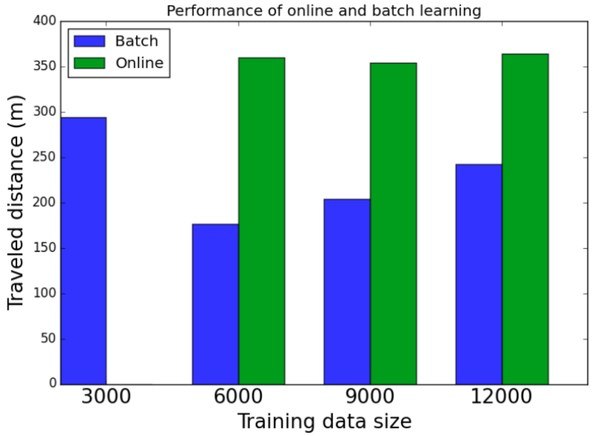
Test result



(a) MPC expert. (b) Batch IL. (c) Online IL.

Experimental Results — Quantitative

- Online IL outperforms Batch IL
- Online IL performance monotonically improves as more training data is involved
- Achieve similar performance (speed) to the expert



Distance traveled without crashing

Policy	Avg. speed	Top speed	Training data	Completion ratio	Total loss	Steering/Throttle loss
Expert	6.05 m/s	8.14 m/s	N/A	100 %	0	0
Batch	4.97 m/s	5.51 m/s	3000	100 %	0.108	0.092/0.124
Batch	6.02 m/s	8.18 m/s	6000	51 %	0108	0.162/0.055
Batch	5.79 m/s	7.78 m/s	9000	53 %	0.123	0.193/0.071
Batch	5.95 m/s	8.01 m/s	12000	69 %	0.105	0.125/0.083
Online (1 iter)	6.02 m/s	7.88 m/s	6000	100 %	0.090	0.112/0.067
Online (2 iter)	5.89 m/s	8.02 m/s	9000	100 %	0.075	0.095/0.055
Online (3 iter)	6.07 m/s	8.06 m/s	12000	100 %	0.064	0.073/0.055

Experimental Results

Limitations

- A task-specific expert is indispensable, which means it is not extensively applicable.
- The experimental setting is simple, only on an empty elliptical dirt track.
- Only show the result from an MPC expert, human-guided imitation learning result is not demonstrated.
- The novelty is medium. The key module (online IL) is from Dagger.

Future Work for Paper / Reading

- How can the imitation learning method perform on more complicated tasks? (obstacles, more curves, different weather)
- How is the generalization ability to more unseen tracks?
- Other robot agents and tasks (UAVs, boats, etc.)

Extended Readings

Off-road autonomous driving:

- Drews, Paul, et al. “Aggressive deep driving: Model predictive control with a cnn cost model.” (2017)
- Williams, Grady, et al. "Aggressive driving with model predictive path integral control.” (2016)
- Williams, Grady, et al. "Information theoretic MPC for model-based reinforcement learning.” (2017)

Extended Readings

Imitation Learning and applications:

- Ross, Stéphane, Geoffrey Gordon, and Drew Bagnell. "A reduction of imitation learning and structured prediction to no-regret online learning.” (2011)
- Zhang, Jiakai, and Kyunghyun Cho. "Query-efficient imitation learning for end-to-end autonomous driving.” (2016)
- Ross, Stéphane, et al. "Learning monocular reactive uav control in cluttered natural environments.” (2013)

Summary

Focus on off-Road Autonomous Driving

- An end-to-end framework instead of the traditional optimal control approach
- Imitation learning method

Batch IL v.s. Online IL: online IL is always better in terms of performance

A successful real-world application for imitation learning theory

SOTA on off-road autonomous driving: good performance, lower cost, higher frequency

Thank you!