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

Carlos Cuartas

October 21, 2021

Yunpeng Pan, Ching-An Cheng, Kamil Saigol, Keuntaek Lee, Xinyan Yan, Evangelos A. Theodorou, and Byron Boots

Motivation

The motivation to this paper was driven by the challenge to design a low-cost robot with the ability to perform high frequency decisions while driving at high speeds.

Helps 3 things about autonomous robots:

- Cheaper parts
- Adaptive
- Real world possibilities

Main Problem - Lack of application for Autonomous Robots

A very big problem we have these days is the lack of application to the real word where uncertainties are everywhere.

Controlled Experimentation:

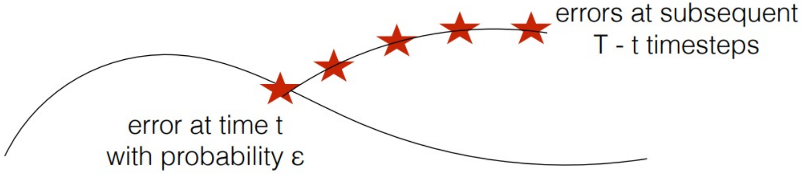
- Lack external validity
- Not applicable to any miniscule changes

Real World Experimentation:

- Applicable to wide variety of situations
- Can ultimately be used in real world scenarios

Things to Know Prior- DAgger

Compounding Errors



- Iterative policy training algorithm
- Reinforcement Learning
- Expert teaches the learner how to recover from past mistakes
- Retrain the main classifier on all states ever encountered by the learner

Simply, DAgger fixes compounding errors that stray from the training data

Related Work / Limitations of Prior Work

Prior Works:

Grady Williams, 2016 “Aggressive driving with model predictive path”

- Expensive hardware
- Only in controlled environments

Paul Drews, 2017 “Aggressive deep driving: Model predictive control with a cnn cost model”

- Uses vision based cost map
- Computationally expensive optimization due to MCP approach

Urs Muller, 2006 “Off-road obstacle avoidance through end-to-end learning”

- Low speed

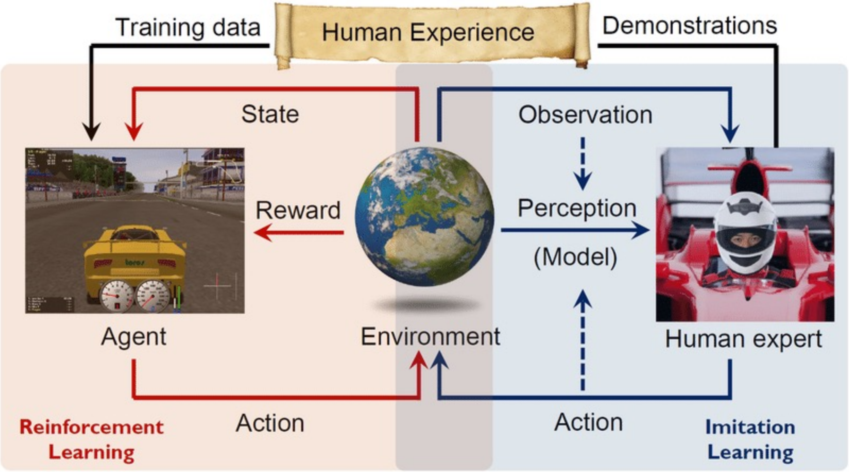
Proposed Approach - Imitation learning

Pros:

- Low chance of damages
- Straight application of knowledge from experts
- Generally better than ordinary RL in real world situations

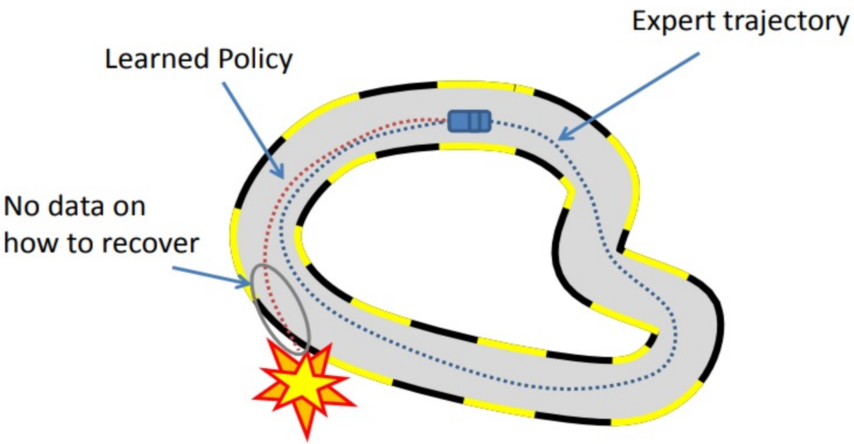
Cons:

- Expert is needed
- Compounding error over time
- When path is strayed data doesn't help

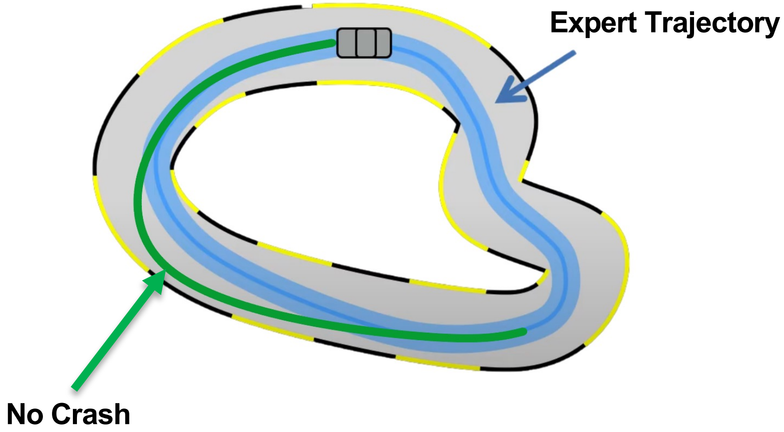


Proposed Approach - DAgger

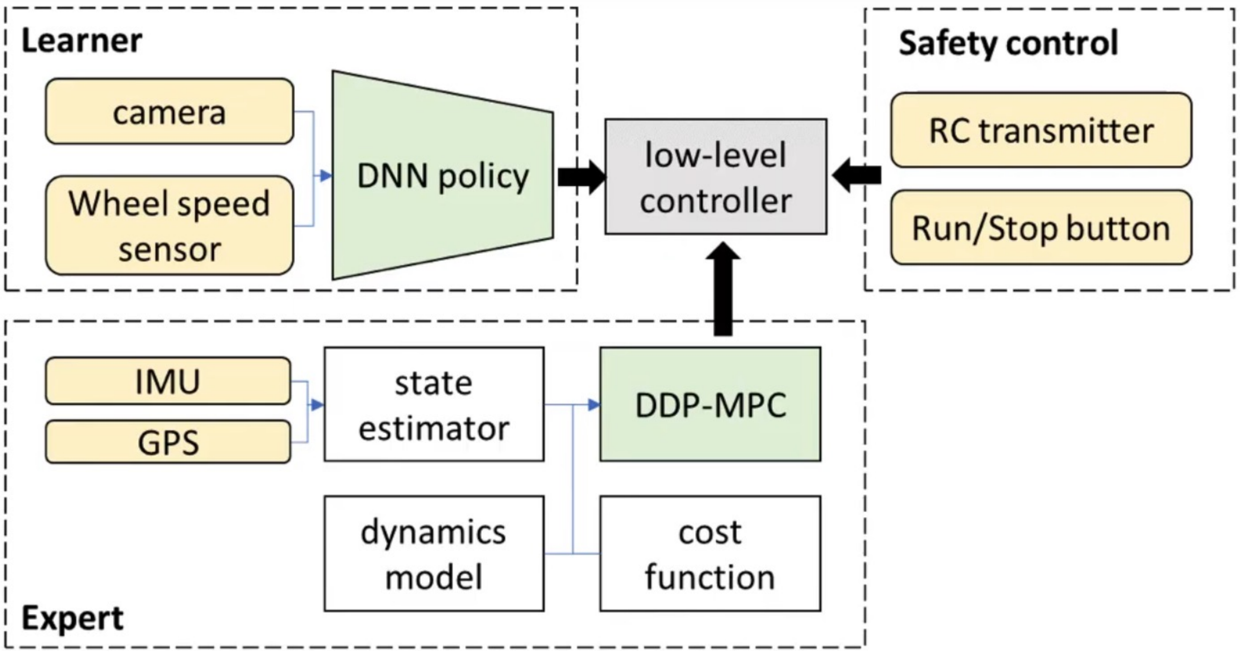
No DAgger:



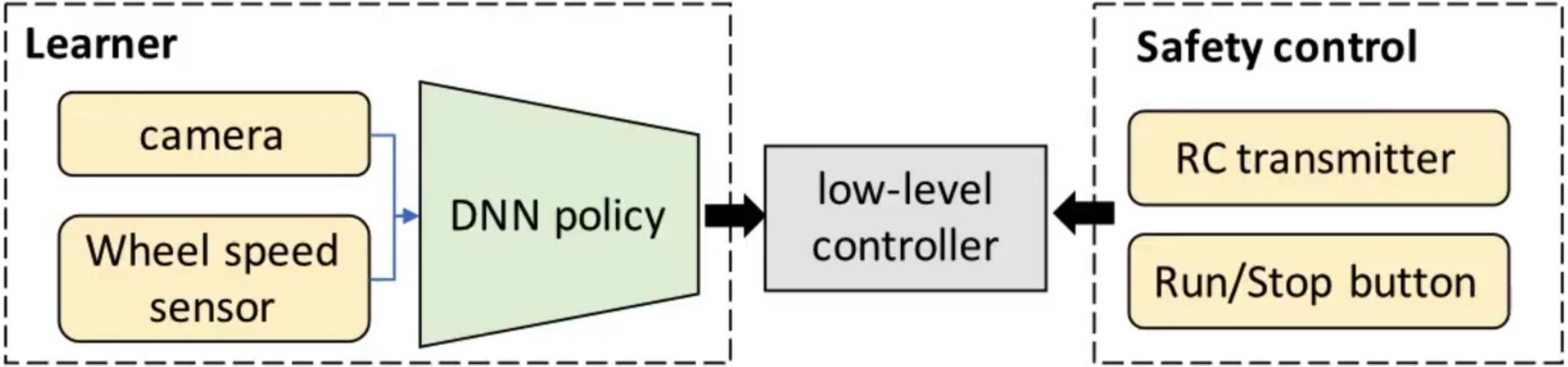
With DAgger:



Proposed Approach – Training Phase



Proposed Approach – Testing Phase



Proposed Approach – Narrowed Objective

Objective:

- Train using Imitation Learning
  - With expert's training samples
- Correct path with Dagger when needed
  - When compounding errors add up
- Be as fast as possible without crashing
  - Hard due to stochastic terrains

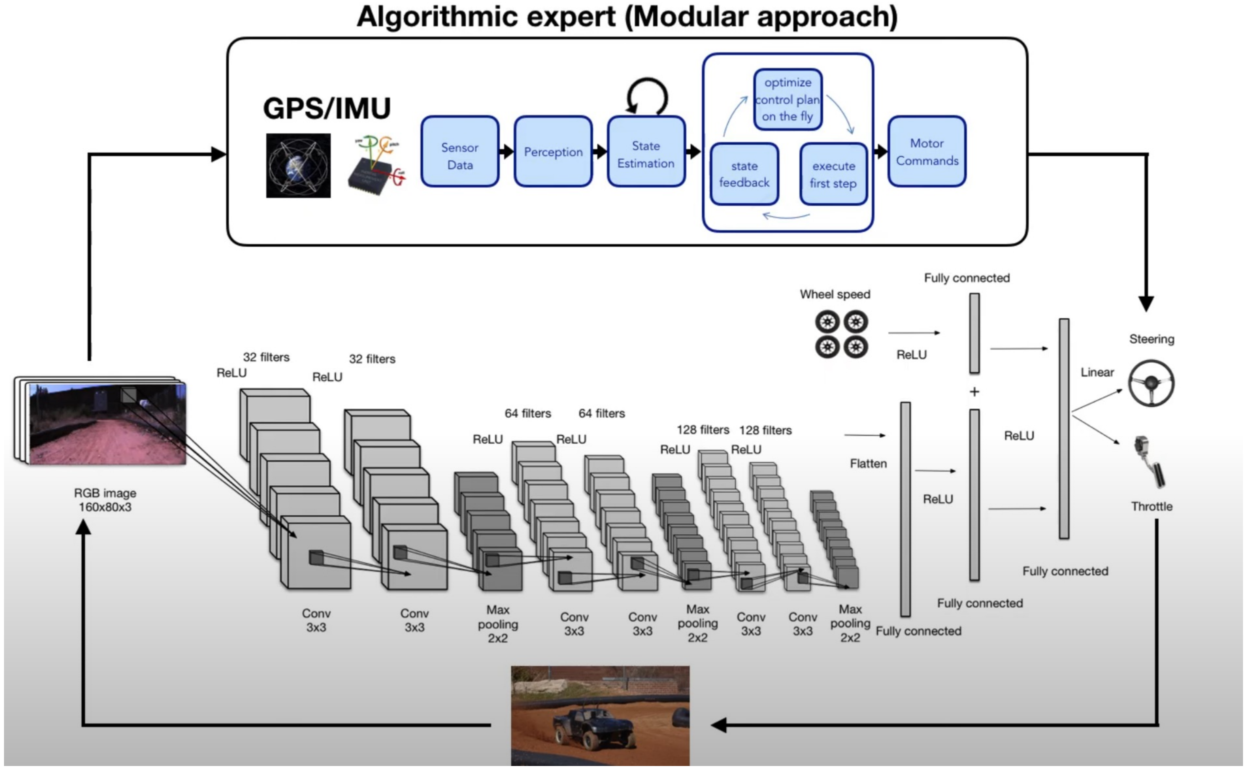
Experimental Setup

- This system was applied to a 1/5 scale Auto Rally car
- The car was equipped with a low cost monocular camera and wheel speed sensors
- The track was a simple off-road track made of dirt
- Desired speeds of about 7 m/s

TABLE I: Comparison of our method to prior work on IL for autonomous driving

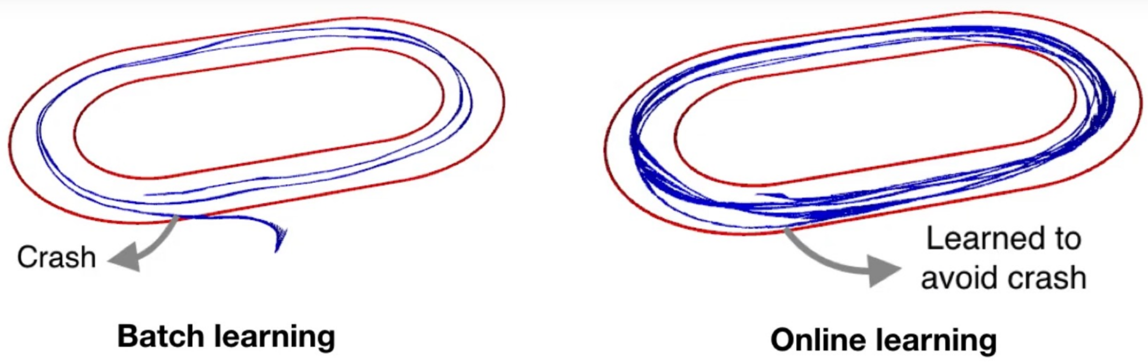
Methods	Tasks	Observations	Action	Algorithm	Expert	Experiment
[1]	On-road low-speed	Single image	Steering	Batch	Human	Real &simulated
[23]	On-road low-speed	Single image & laser	Steering	Batch	Human	Real &simulated
[24]	On-road low-speed	Single image	Steering	Batch	Human	Simulated
[20]	Off-road low-speed	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

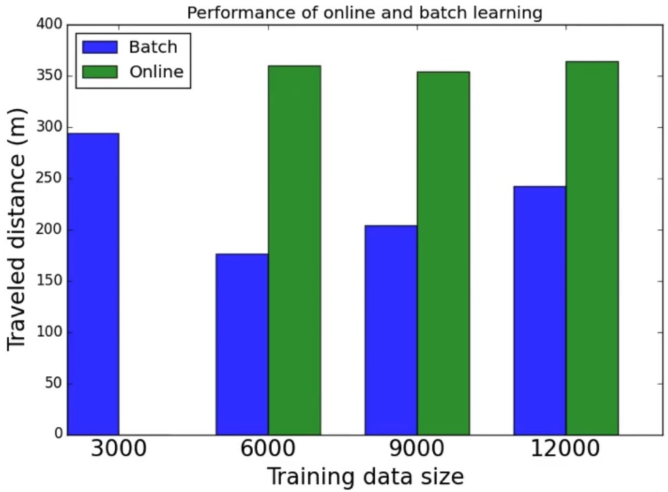
Experimental Setup – Proposed System



Results

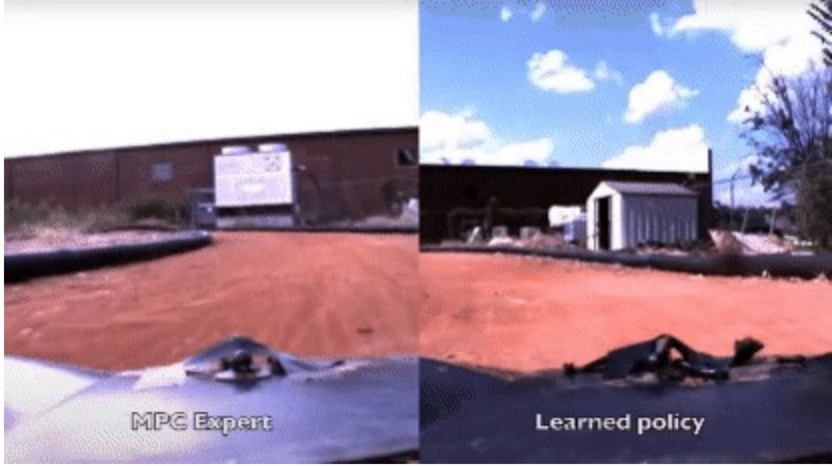
Results


Crash
Learned to
avoid crash
Batch learning
Online learning

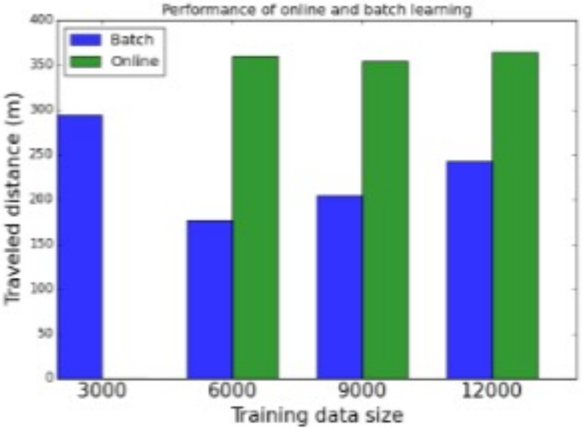

Performance of online and batch learning
Batch
Online
Traveled distance (m)
400
350
300
250
200
150
100
50
0
3000
6000
9000
12000
Training data size

Discussion of Results

- Able to achieve same speeds as MPC expert
- Online data (DAgger) out preformed the batch data
- Found that online IL always improves as more data is gathered, this opposes other research papers on IL



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



Critique / Limitations

- Impressive due to the implementation of DAgger and pushing the state of the art but the novelty is medium at best since many other papers have covered this topic
- Only truly applicable in real life when an expert is there to provide data samples
- This can still be matched or passed with other methods using very high sampling but the risk of crashing is higher
- Even with Imitation learning, it is still very possible for a crash to happen

Future Work for Paper / Reading

- Drones
- Experimentation on many tracks
- Multiple expert applications
- Application of new sensors like depth cameras or Gyroscopes
- Assuring there will be no crashes even with new or unknown entities on track

Extended Readings 1

Grady Williams, Paul Drews, Brian Goldfain, James M Rehg, and Evangelos A Theodorou. **“Aggressive driving with model predictive path integral control.”** 2016

Paul Drews, Grady Williams, Brian Goldfain, Evangelos A Theodorou, and James M Rehg. **“Aggressive deep driving: Model predictive control with a cnn cost model.”** 2017

Urs Muller, Jan Ben, Eric Cosatto, Beat Flepp, and Yann L Cun. “**Off-road obstacle avoidance through end-to-end learning.**” 2006

Stéphane Ross, Geoffrey J. Gordon, J. Andrew Bagnell. **“A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning”** 2010

Extended Readings 2

Dean A Pomerleau. **“Alvinn: An autonomous land vehicle in a neural network”** 1989

Viktor Rausch, Andreas Hansen, Eugen Solowjow, Chang Liu, Edwin Kreuzer, and J. Karl Hedrick. **“Learning a deep neural net policy for end-to-end control of autonomous vehicles.”** 2017

Jeff Michels, Ashutosh Saxena, and Andrew Y Ng. **“High speed obstacle avoidance using monocular vision and reinforcement learning.”** 2005.

Summary

- Addresses the lack of application Autonomous Robots to the real word where uncertainties are everywhere.
- The use of stochastic terrains for intentionally driving the robot off track similar to the real world
- Discuss the shortcomings of other robots like low speed, only simulated, or not off road
- This paper uses Imitation Learning, the Dagger method to keep it on track, and DNN Control Policy
- Pushed the bounds of state of the art for End-to-End Autonomous Driving using IL