Reliable Operation of Machine Learning Models in Autonomous Driving Systems

Lishan Yang

Incoming Assistant Professor in CS@GMU

lyang11@email.wm.edu

Evgenia Smirni (Professor)

Students:

Anna Schmedding

Philip Schowitz

Homa Alemzadeh (Professor)

Students:

Xugui Zhou

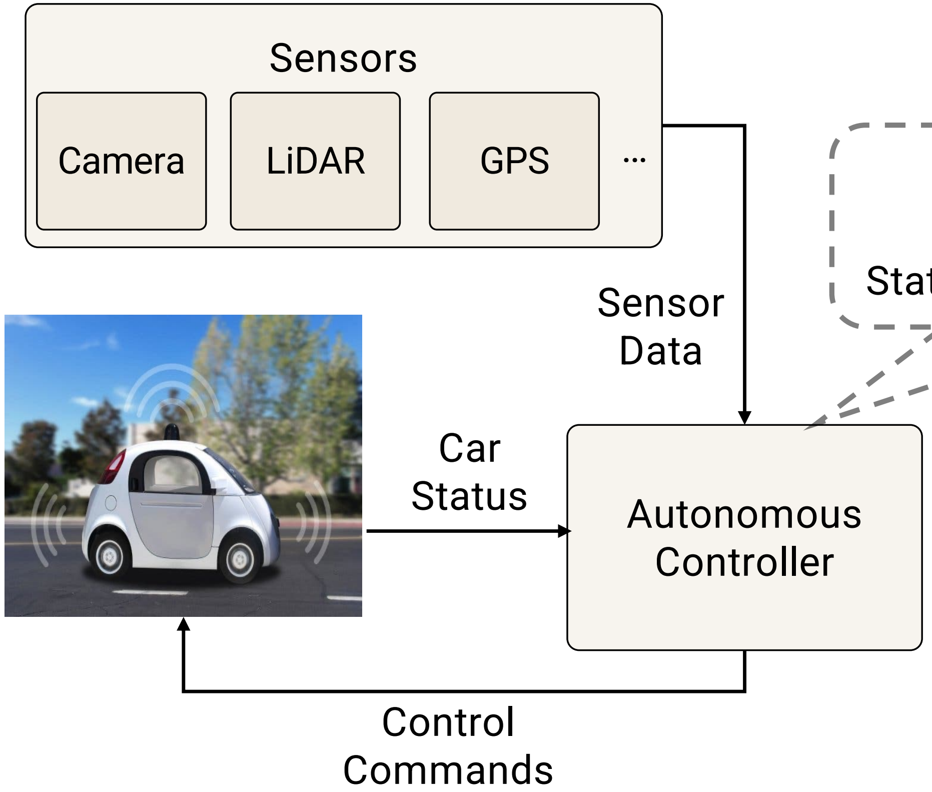
Haotian Ren

WILLIAM & MARY
CHARTERED 1693

GEORGE MASON UNIVERSITY

UNIVERSITY of VIRGINIA

Autonomous Driving Systems



Sensors

Camera

LiDAR

GPS

...

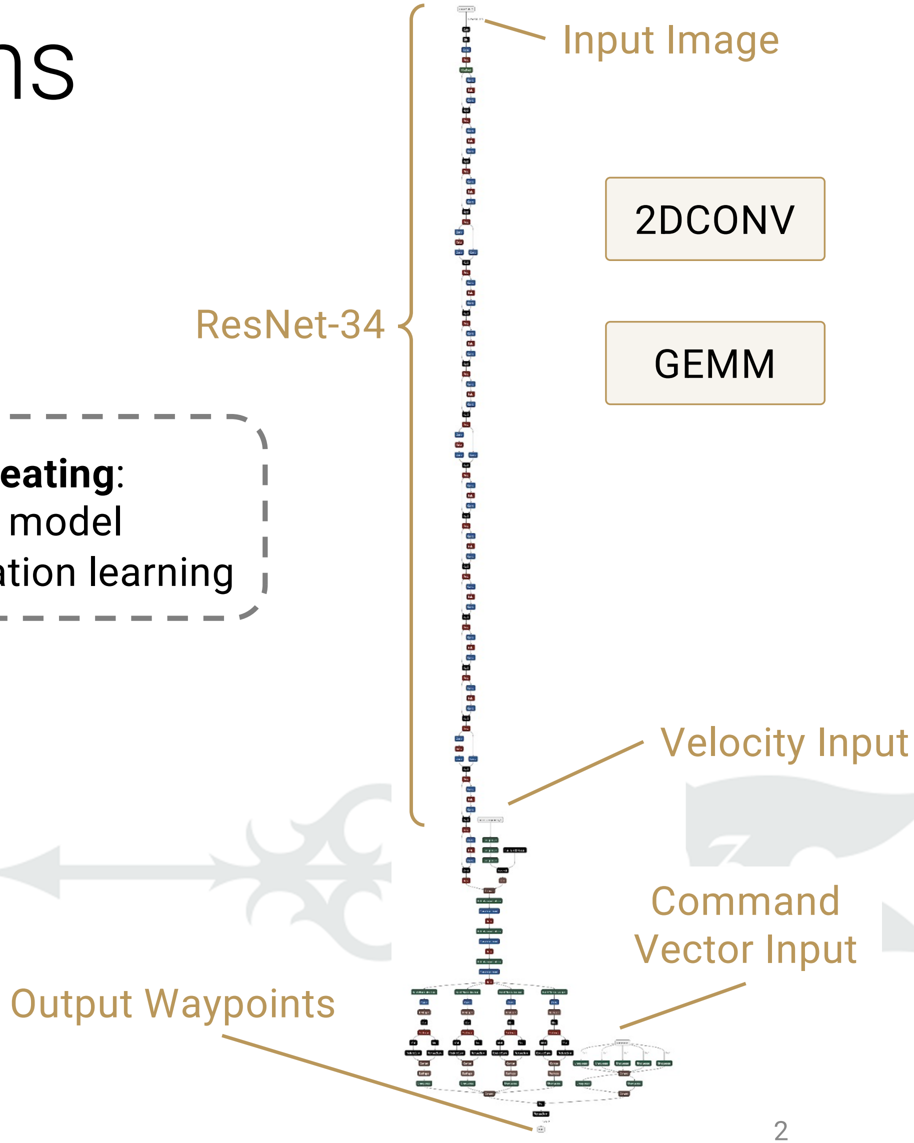
Sensor Data

Car Status

Autonomous Controller

Control Commands

LearningByCheating:
An end-to-end model
State-of-the-art imitation learning



Input Image

ResNet-34

2DCONV

GEMM

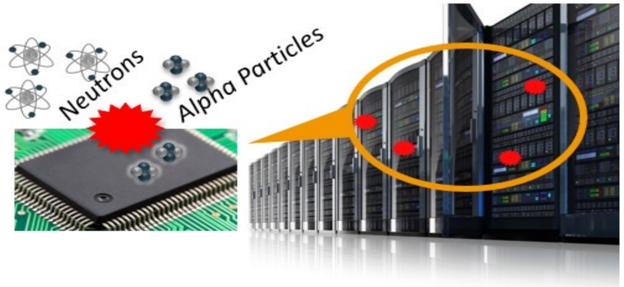
Velocity Input

Command Vector Input

Output Waypoints

Chen, Dian, et al. "Learning by cheating." Conference on Robot Learning. PMLR, 2020.

Reliability in Autonomous Driving Systems



- Soft errors:
  - The most commonly observed errors
  - Source: High-energy radioactive particles (i.e., cosmic rays)
  - Bit flips

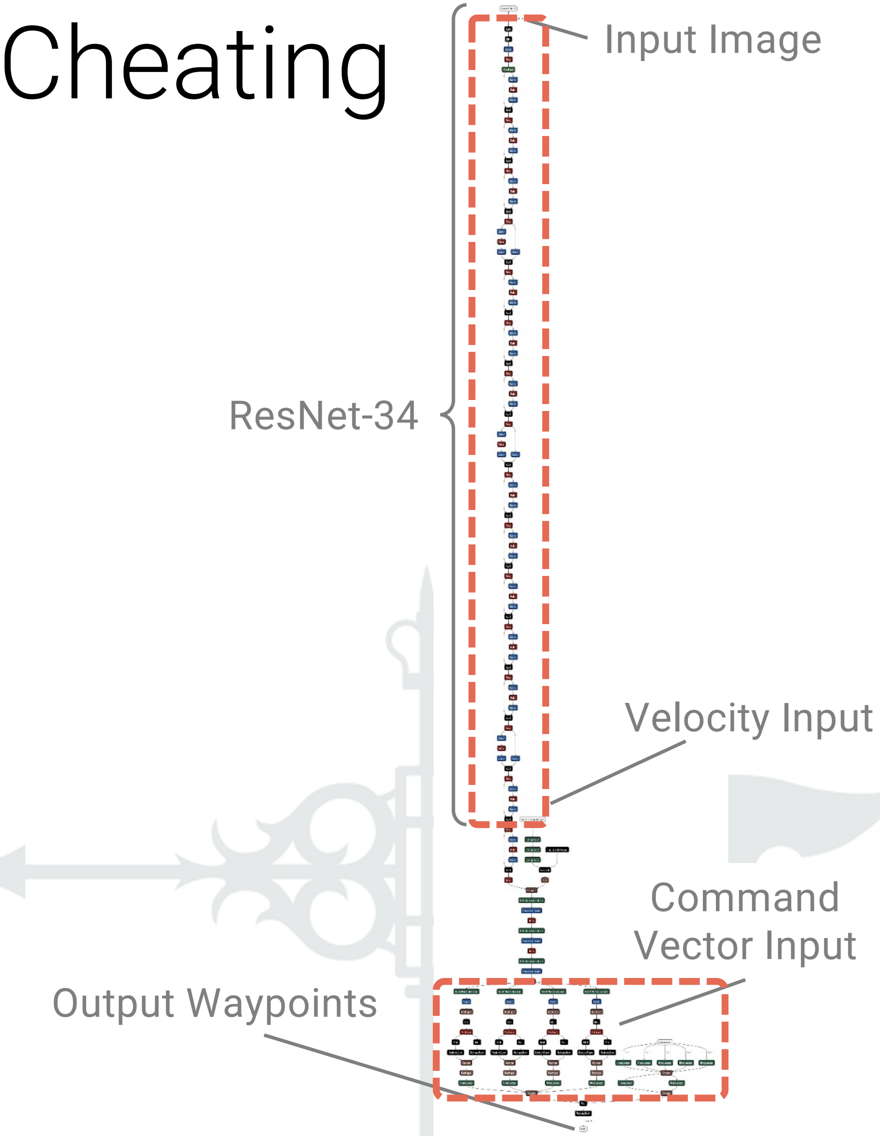
- ***Safe***
- ***Hazards***
- ***Crashes***

http://global.atpinc.com/Memory-insider/what-is-soft-error-detection-sram-emmc

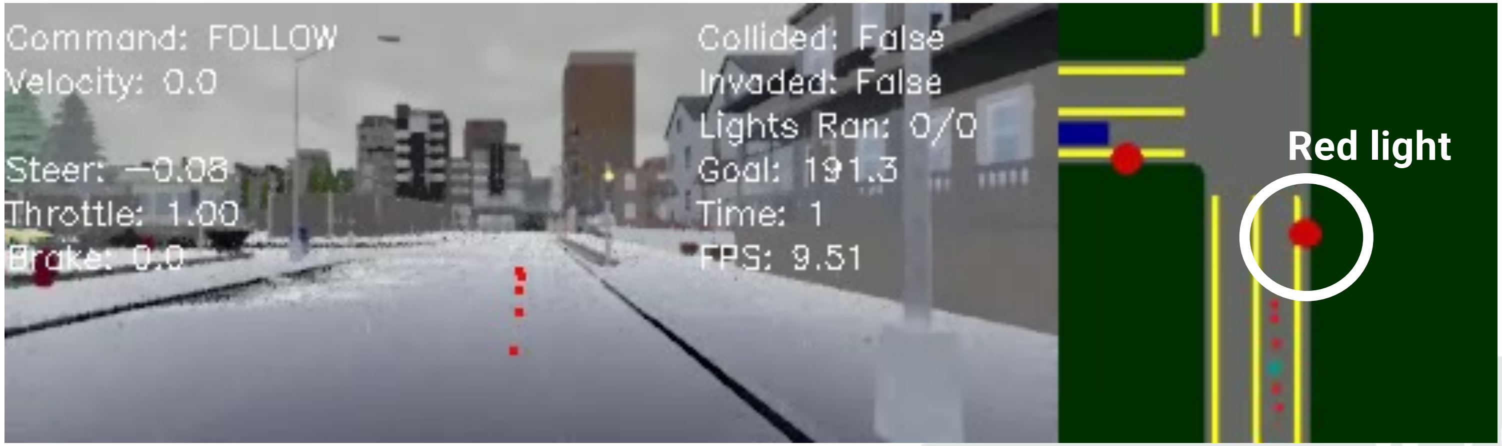
Fault Injection in Learning by Cheating

- Double-bit flip in one weight
  - Convolution layers in ResNet-34
  - Convolution layers in branches at the end

- Outcomes:
  - Safe
  - Immediate crash
    - Turns sharply and drives off road
  - End branches: strange behaviors
    - Turns corners too wide
    - Lane invasions and swerving
    - Runs red lights
    - Crash



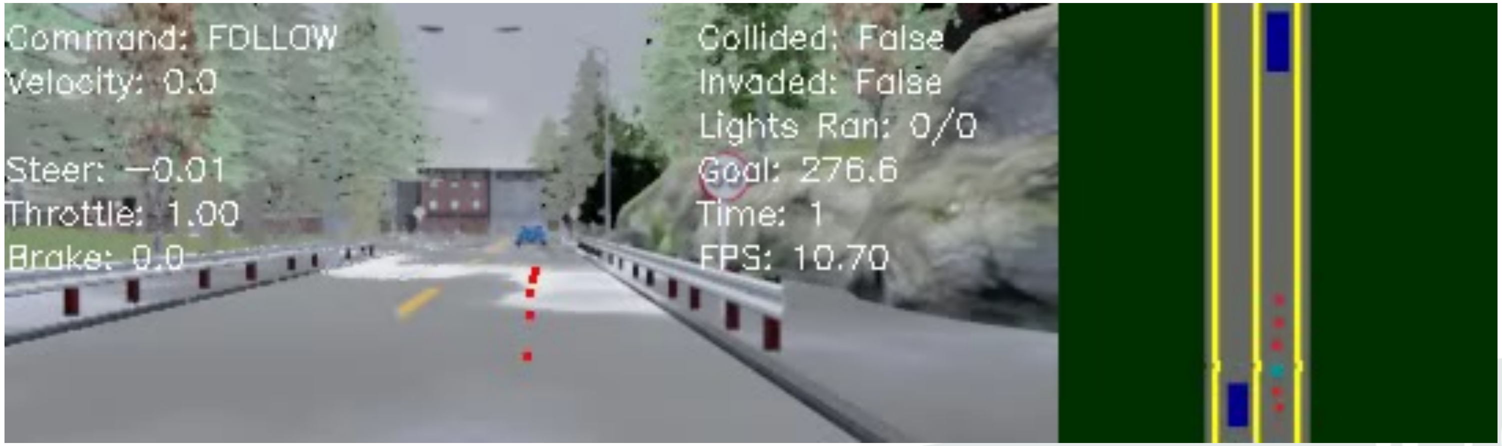
Red-Light Violation After Fault Injection



Command: FOLLOW
Velocity: 0.0
Steer: −0.08
Throttle: 1.00
Brake: 0.0

Collided: False
Invaded: False
Lights Ran: 0/0
Goal: 191.3
Time: 1
FPS: 9.51

Crash After Fault Injection



```
Command: FOLLOW
Velocity: 0.0

Steer: −0.01
Throttle: 1.00
Brake: 0.0
```

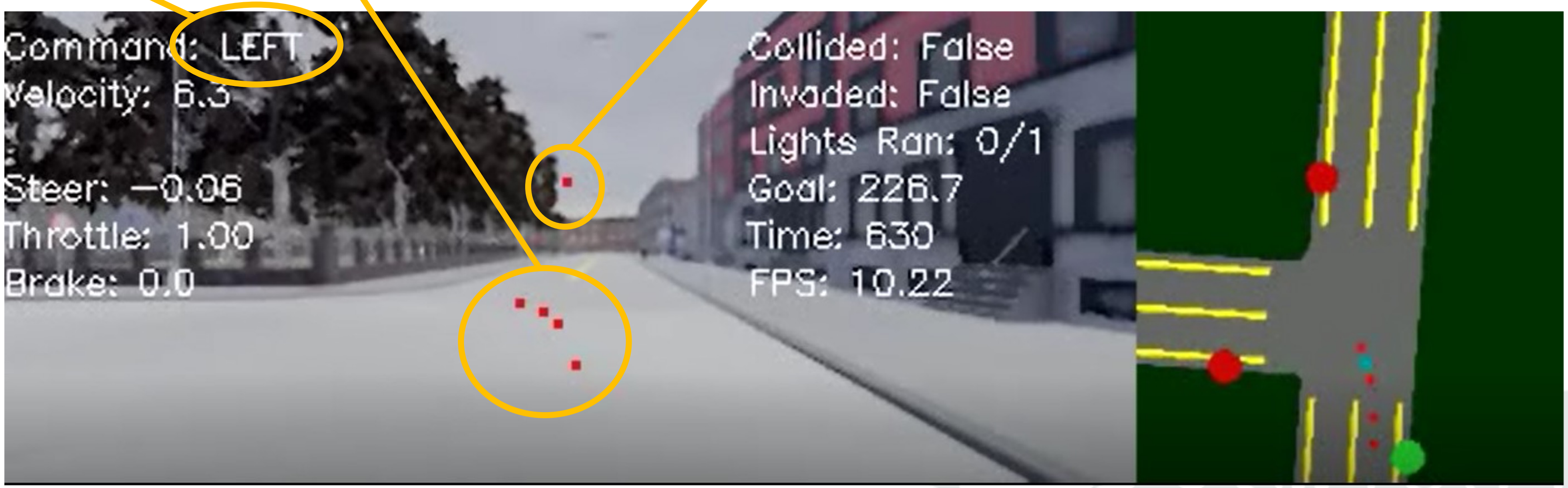
```
Collided: False
Invaded: False
Lights Ran: 0/0
Goal: 276.6
Time: 1
FPS: 10.70
```

What Causes the Crash?

Turn-left command is vulnerable

Reasonable waypoints

Clear incorrect waypoint



- One waypoint deviates significantly
- Car turns too widely to attempt to fit the curve

Effects of Environmental Conditions

- Weather
  - Bad weather (e.g., rain): more vulnerable
  - More red-light violation

- Sharp turns in roads
  - Higher chance of lane invasions and crashes

- Pedestrians and other vehicles
  - Lane invasions more likely to hit pedestrians and other vehicles

Ongoing Work

- A strong characterization
  - Systematic fault injection experiments
- Analysis of important weights for Learning By Cheating
  - Any proxy?
- Low-overhead Protection

Reliability Autonomous Driving Systems

- Bit Flips → Hazards & Collisions
- We need protection!

Other Ways of Fault Injection?

Hardware-Level Protection?

Other ADS designs?

Deploy on an Autonomous Vehicle?

Security?

Reliable Operation of Machine Learning Models in Autonomous Driving Systems

Lishan Yang

Incoming Assistant Professor in CS@GMU

lyang11@email.wm.edu

Evgenia Smirni (Professor)

Students:

- Anna Schmedding
- Philip Schowitz

Thank you

Homa Alemzadeh (Professor)

Students:

- Xugui Zhou
- Haotian Ren

WILLIAM & MARY

CHARTERED 1693

GEORGE MASON UNIVERSITY

UNIVERSITY of VIRGINIA