Lecture 4: Safety III

Professor Katie Driggs-Campbell

January 25, 2024

ECE484: Principles of Safe Autonomy

Videos courtesy of Tianchen Ji

Administrivia

- Lab starts this week – will introduce MP0

Core Idea of Inductive Invariants

Adding more information

$\text{timer} \coloneqq 0$

if $x_2 - x_1 < d_s$

    if $v_1 > a_b$

        $v_1 \coloneqq v_1 - a_b$

        $\text{timer} \coloneqq \text{timer} + 1$

    else $v_1 = 0$

else $v_1 \coloneqq v_1$

$x_1 \coloneqq x_1 + v_1$

$x_2 \coloneqq x_2 + v_2$

Three Cases to Consider: (1)

Three Cases to Consider: (2)

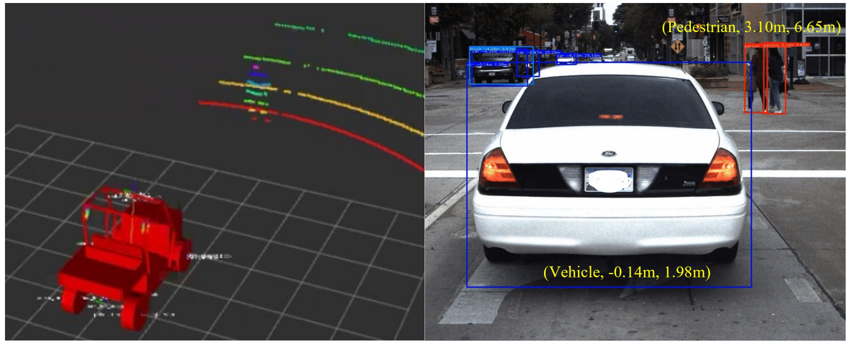
Three Cases to Consider: (3)

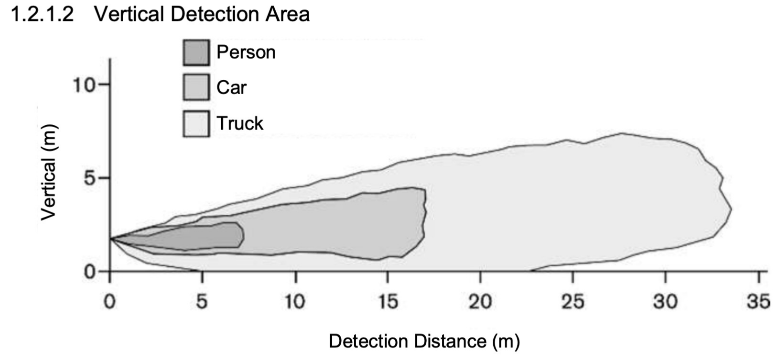
Showing Safety with a Timer

- Goal: show $x_2 - x_1 > 0$
- Maximum distance traveled by car 1 after detection:

Baked-in Assumptions (1)

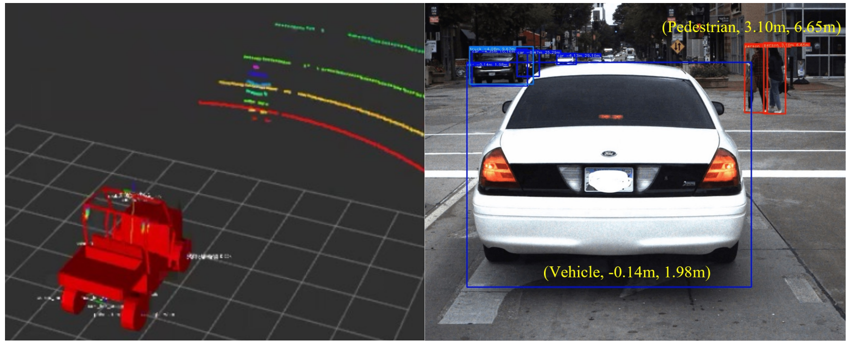
- Perception.
  - Sensor detects obstacle **iff** distance $d \leq D_{sense}$
  - How to model vision errors?



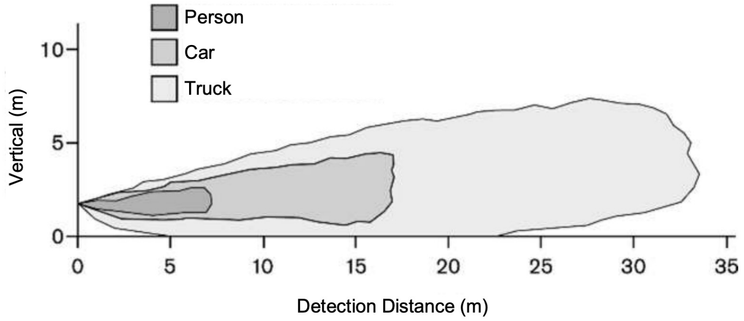


Baked-in Assumptions (1)

- Perception.
  - Sensor detects obstacle **iff** distance $d \leq D_{sense}$
  - How to model vision errors?
- Pedestrian Behaviors.
  - Pedestrian is assumed to be moving with constant velocity from initial position

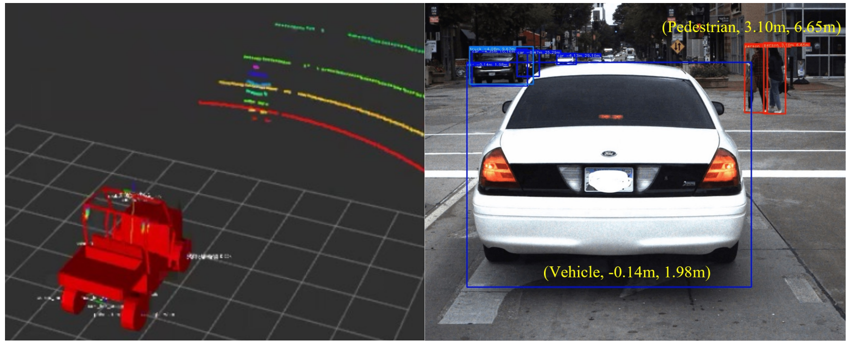


1.2.1.2 Vertical Detection Area

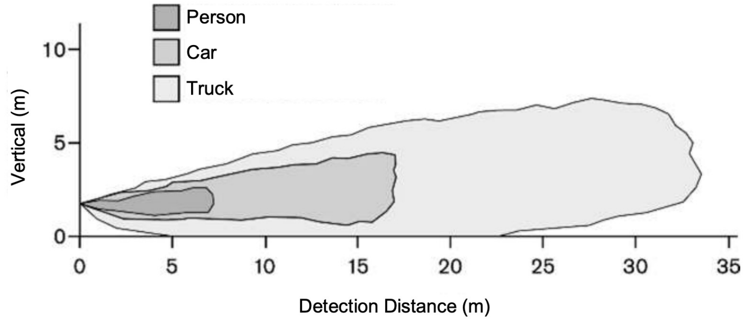


Baked-in Assumptions (1)

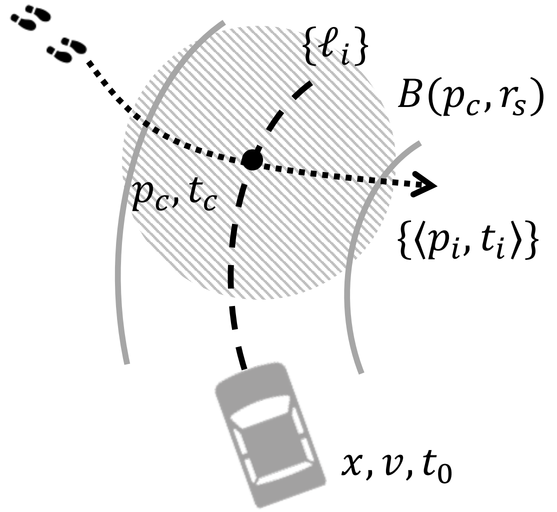
- Perception.
  - Sensor detects obstacle **iff** distance $d \leq D_{sense}$
  - How to model vision errors?
- Pedestrian Behaviors.
  - Pedestrian is assumed to be moving with constant velocity from initial position
- No sensing-computation-actuation delay.
  - The time step in which $d \leq D_{sense}$ is true is exactly when the velocity starts to decrease



1.2.1.2 Vertical Detection Area

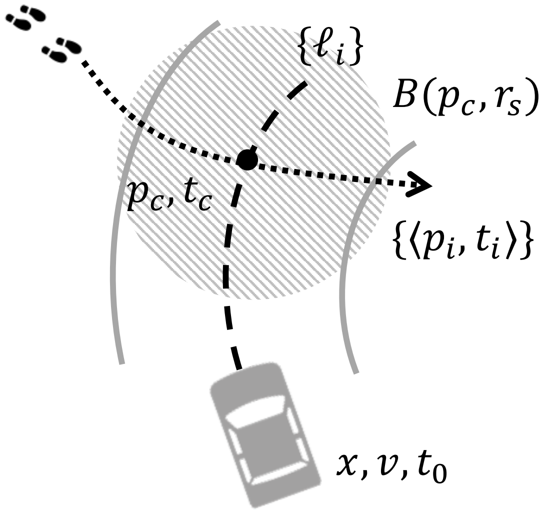


Baked-in Assumptions (2)



- Mechanical or Dynamical assumptions
  - Vehicle and pedestrian moving in 1-D lane.
  - Does not go backwards.
  - Perfect discrete kinematic model for velocity and acceleration.

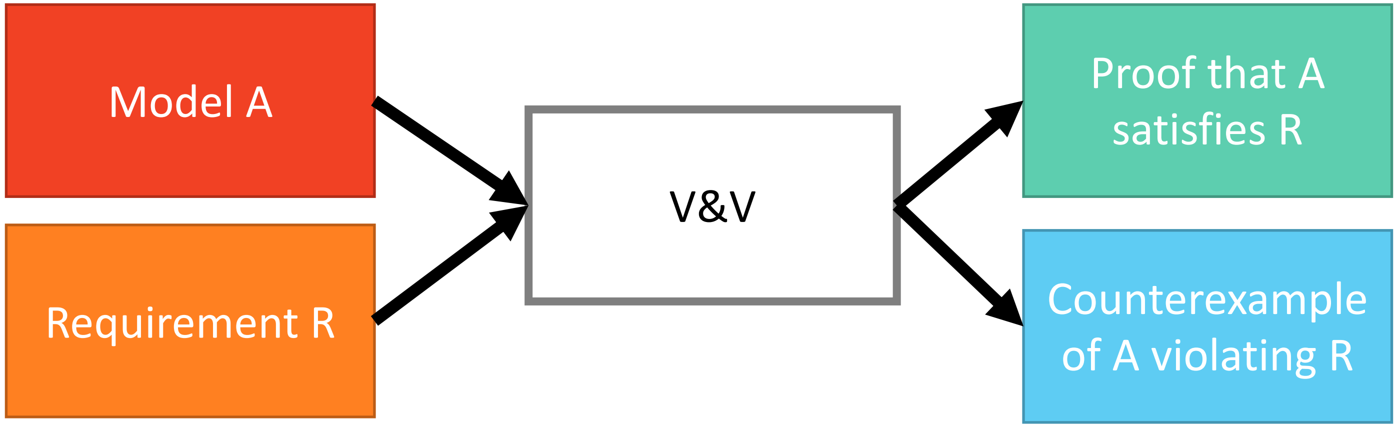
Baked-in Assumptions (2)



- Mechanical or Dynamical assumptions
  - Vehicle and pedestrian moving in 1-D lane.
  - Does not go backwards.
  - Perfect discrete kinematic model for velocity and acceleration.
- Nature of time
  - Discrete steps. Each execution of the above function models advancement of time by 1 step. If 1 step = 1 second, $x_1(t+1) = x_1(t) + v_1(t).1$
  - Atomic steps. 1 step = complete (atomic) execution of the program.
    - We cannot directly talk about the states visited after partial execution of program

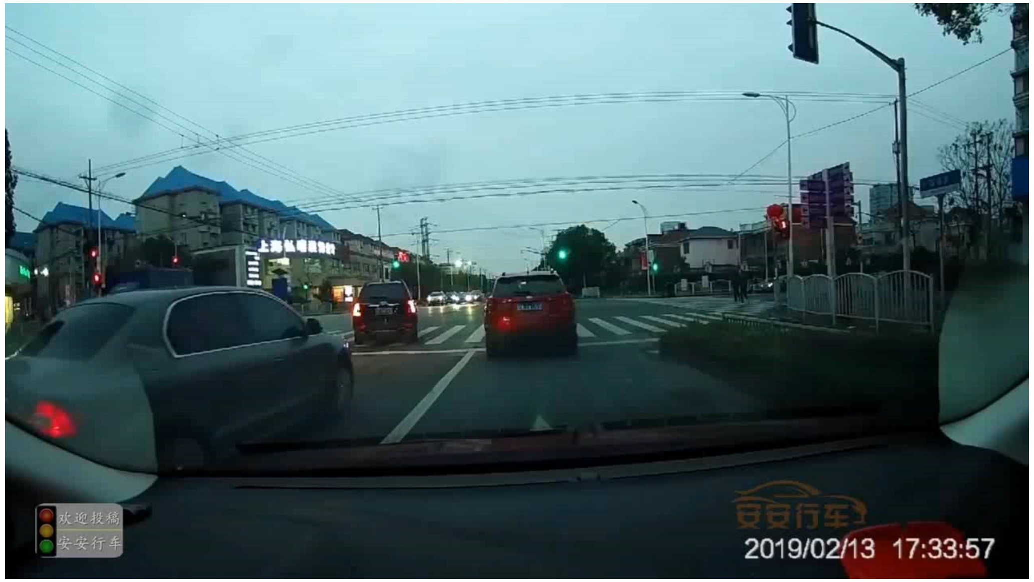
Remarks and Takeaway

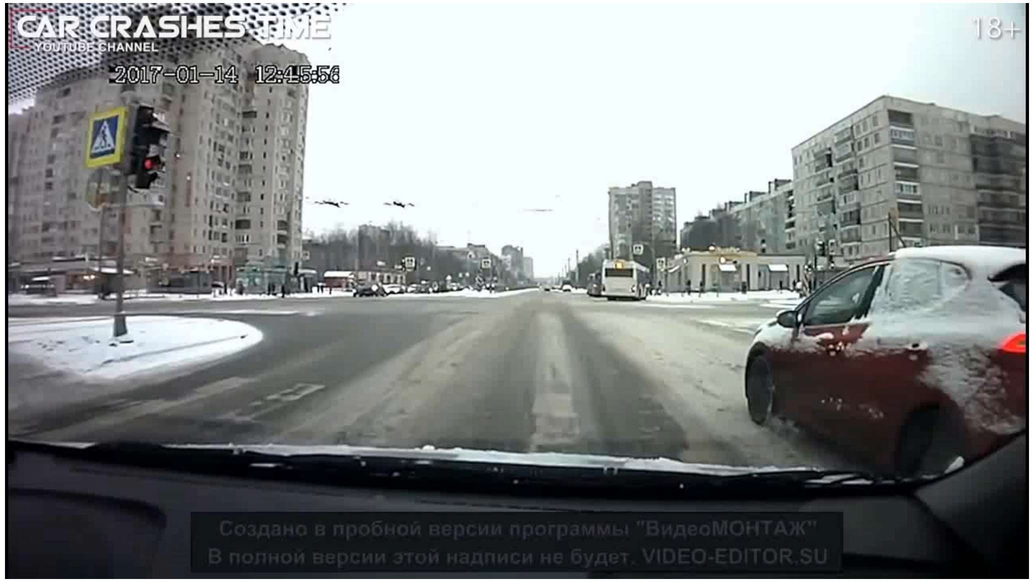
- The proof by induction shows a property of ***all behaviors of our model***
- The proof is conceptually simple, but can quickly get tedious and error prone
  - Verification and Validation tools like Z3, Dafny, PVS, CoQ, AST, MC2, automate this



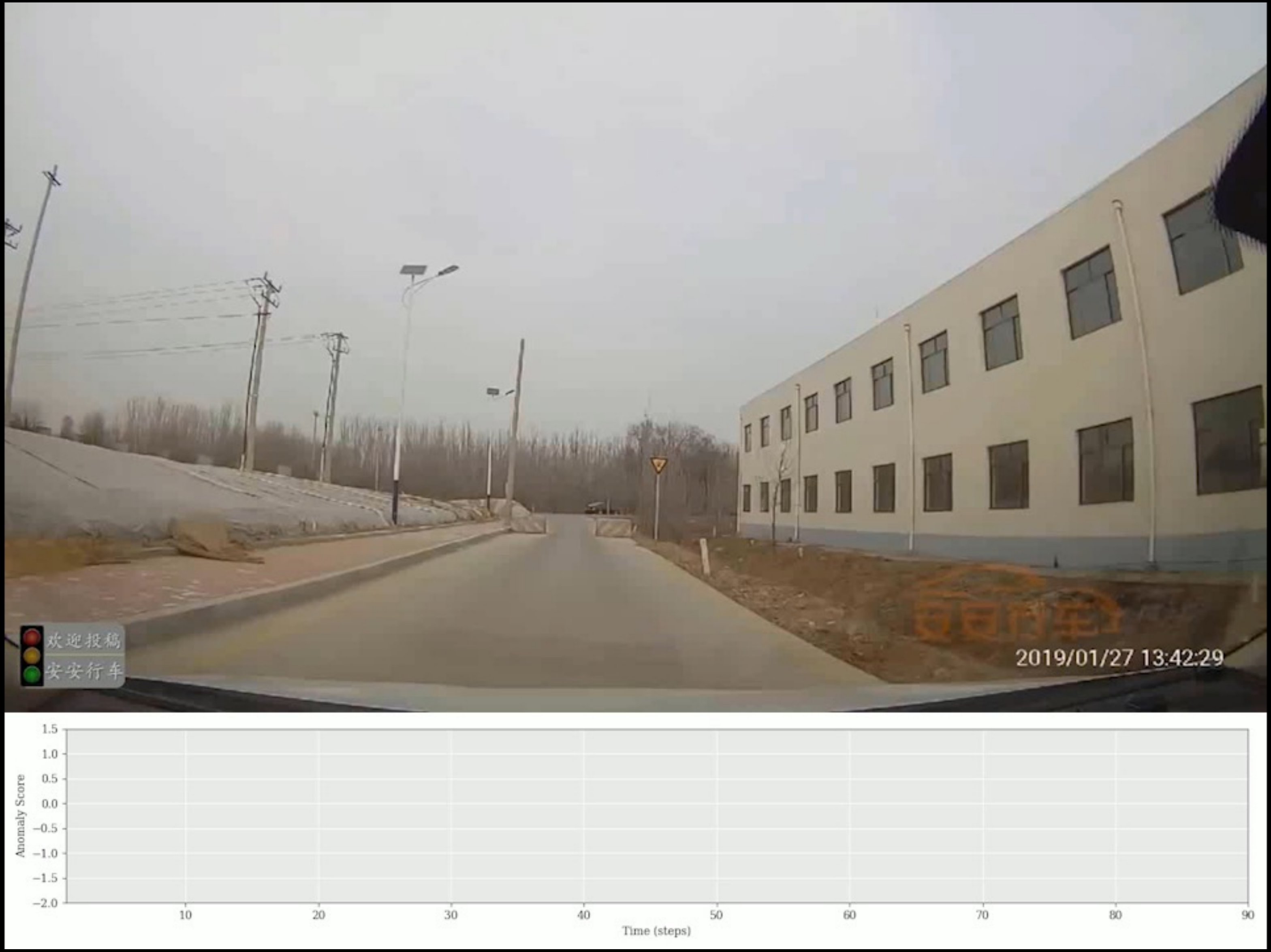
Rare Events and Safety Proxies

Anomalies in Driving Scenes

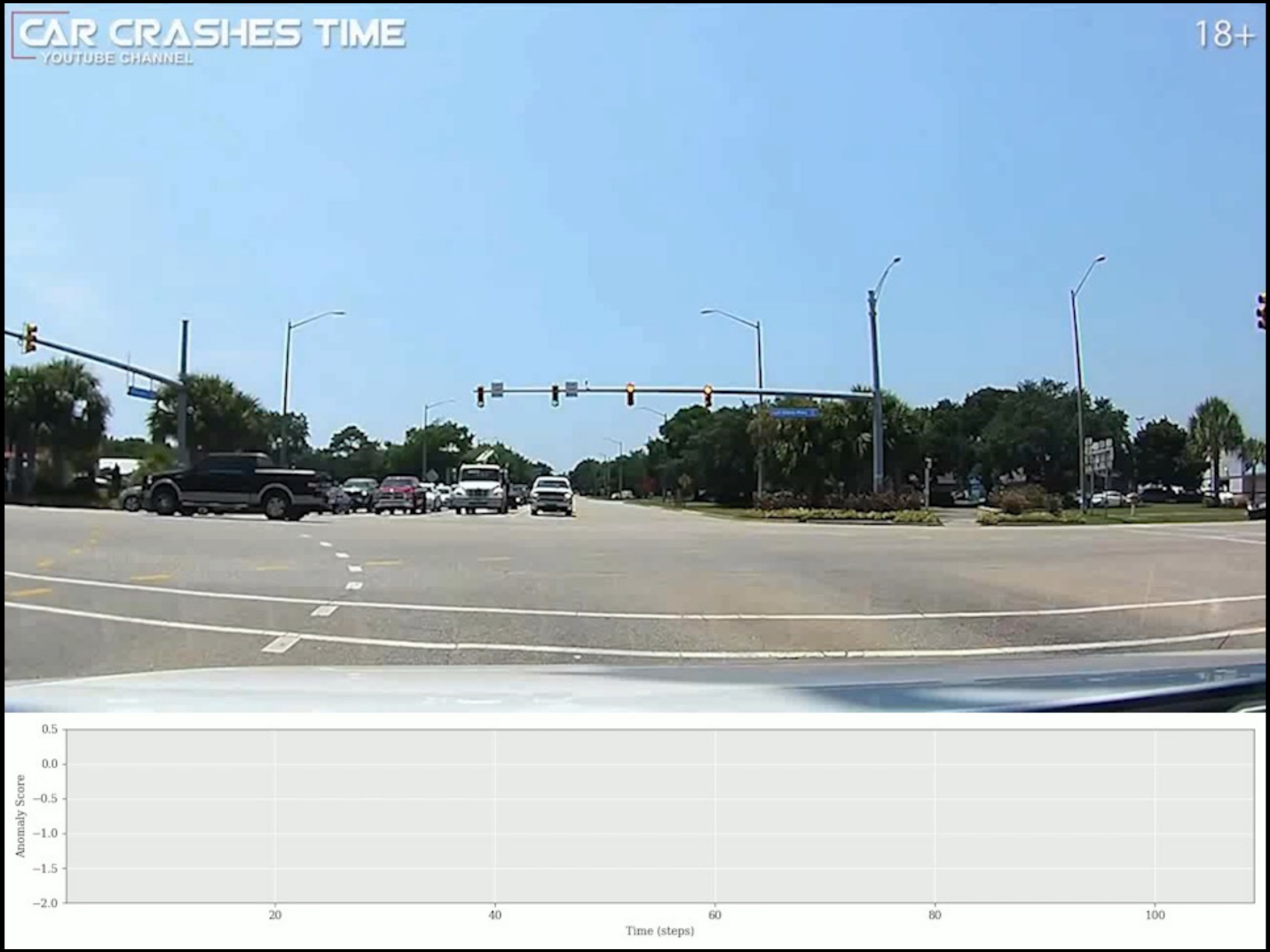




On-Road Anomaly Detection



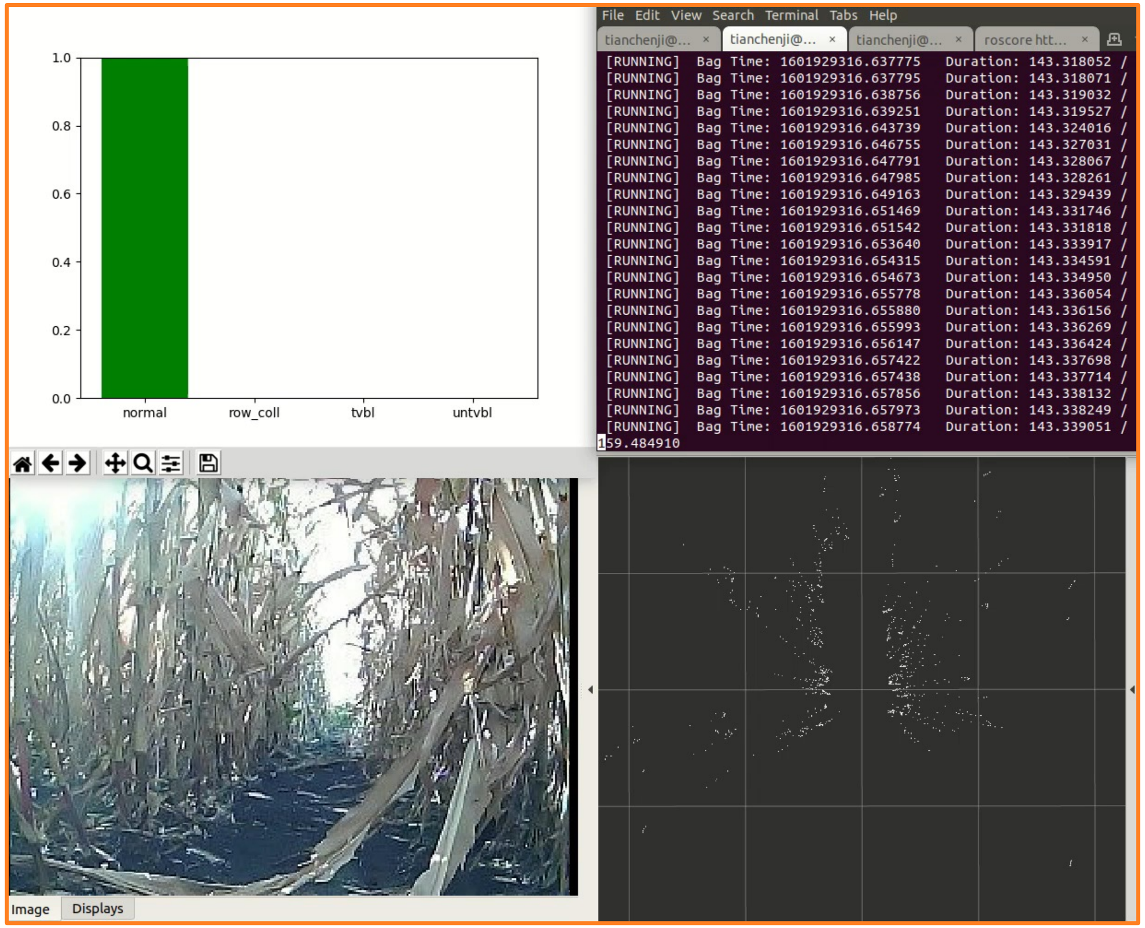
On-Road Anomaly Detection



Anomalies in Field Environments

Reactive Anomaly Detection

Reactive Anomaly Detection



Proactive Anomaly Detection

Summary

- Invariant trick can give a shortcut for proving safety ☺
  - The invariant I may contain important information about conserved quantities and may also tell us why the system is safe
  - However, often requires guessing and checking and a lot of engineering effort
- Online Monitoring is another key component to safe systems
  - Anomaly detection is a reasonably proxy for safety, if you don’t mind false positives
- Next week: starting perception