Designing Safe Autonomous Cyber-Physical Systems

Shreyas Ramakrishna (shreyas.ramakrishna@vanderbilt.edu)
Graduate Research Assistant @ scopelab
Institute For Software Integrated Systems, Vanderbilt University
https://www.shreyasramakrishna.com/

Tel (615) 343-7472 Fax (615) 343-7440
1025 16th Avenue South|Nashville, TN 37212
www.isis.vanderbilt.edu

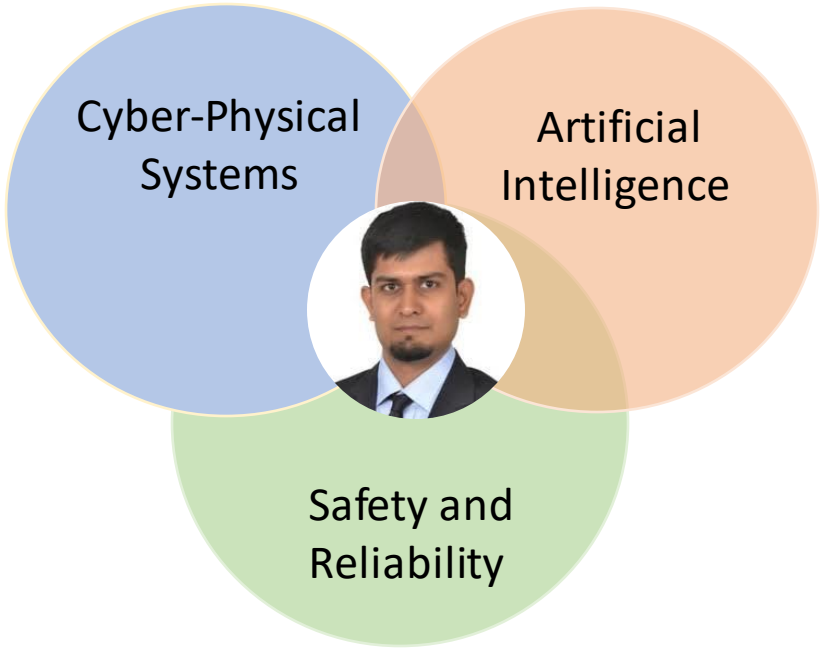
About Me

Education

- Currently, A fifth year Ph.D. candidate @ Institute For Software Integrated Systems, Vanderbilt University, working @ scope lab with Professor Abhishek Dubey for DARPA's Assured Autonomy Project.
- Masters in Electrical Engineering from Technical University Kaiserslautern (Germany), with Master Thesis @ Department of Cyber Physical-Systems.
- Bachelors in Electronics and Communication Engineering from Visvesvaraya Technological University (VTU), India.

Work Experience

- Embedded Design Engineer @ Apsis Solutions, Bangalore, India.

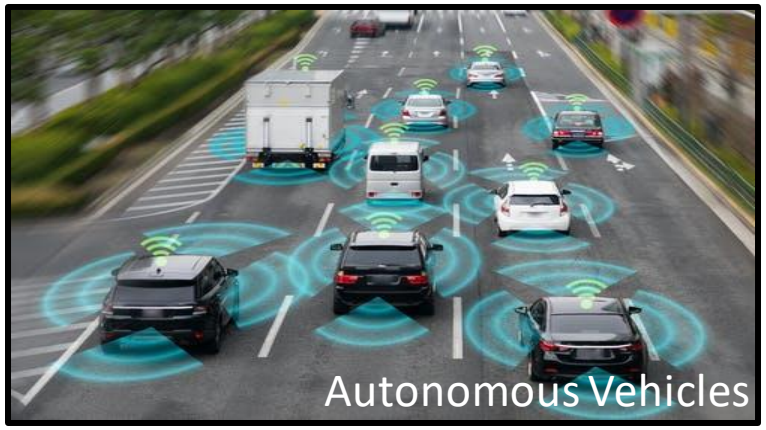


Research Interests

- Cyber Physical Systems
- Artificial Intelligence
- Risk and Reliability

Cyber-Physical Systems

CPS = Computation + Communication + Control

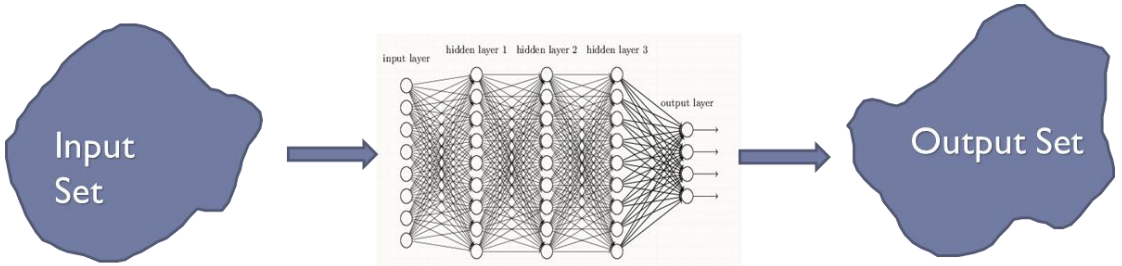


It is important to build CPS that

- *anticipate change: uncertain environments, faults, updates.*
- *exhibit* resilience*: they survive and adapt to faults, while being dependably functional.*
- *are safe: it is important that the systems under operation are shown to be safe.*

Autonomous Cyber-Physical Systems

- Not 'designed' but 'trained' (on data)
  - Design: Architecture + Training method
  - Training:
    - Supervised learning
    - Unsupervised learning
    - Reinforcement learning
- AI-CPS has been widely used in several real world applications.



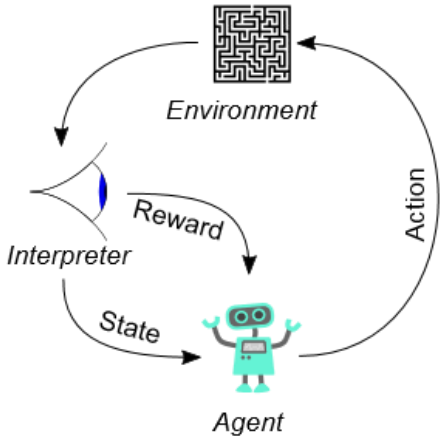
Supervised learning

Factory bots

Hospital bots

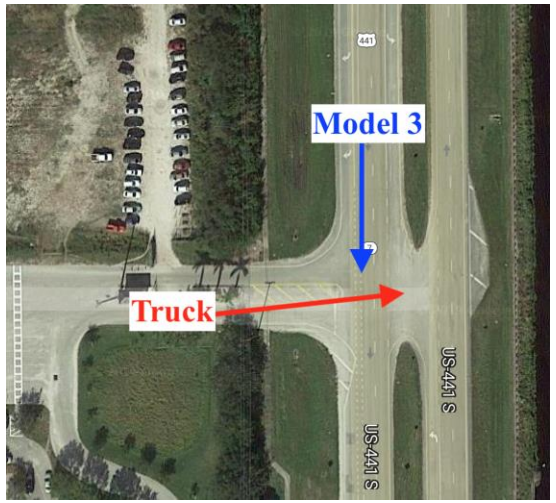
Delivery drones

Autonomous Cars

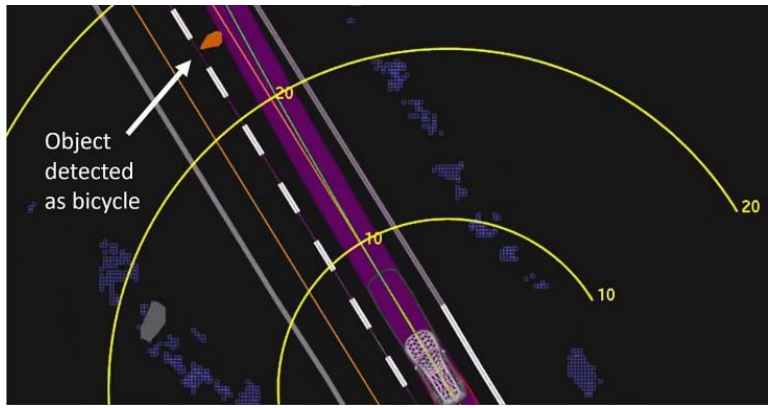


Reinforcement learning

Problem of Safety



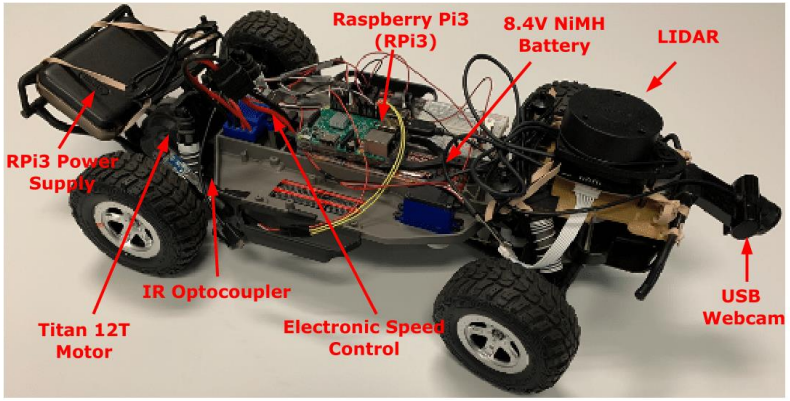
Tesla's autopilot crash



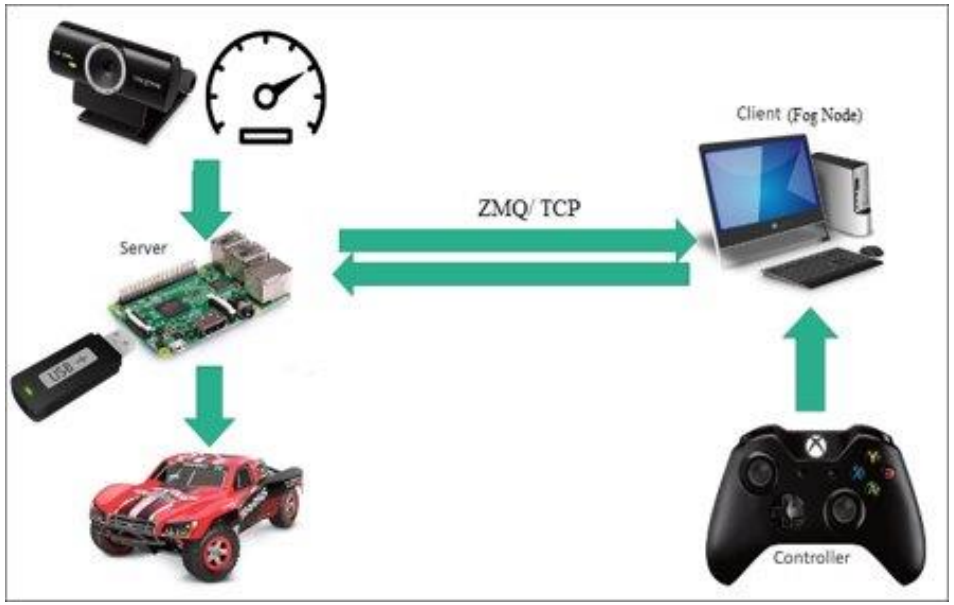
Uber's self driving accident

"self-driving cars are the natural extension of active safety and obviously something we should do."
-elon musk

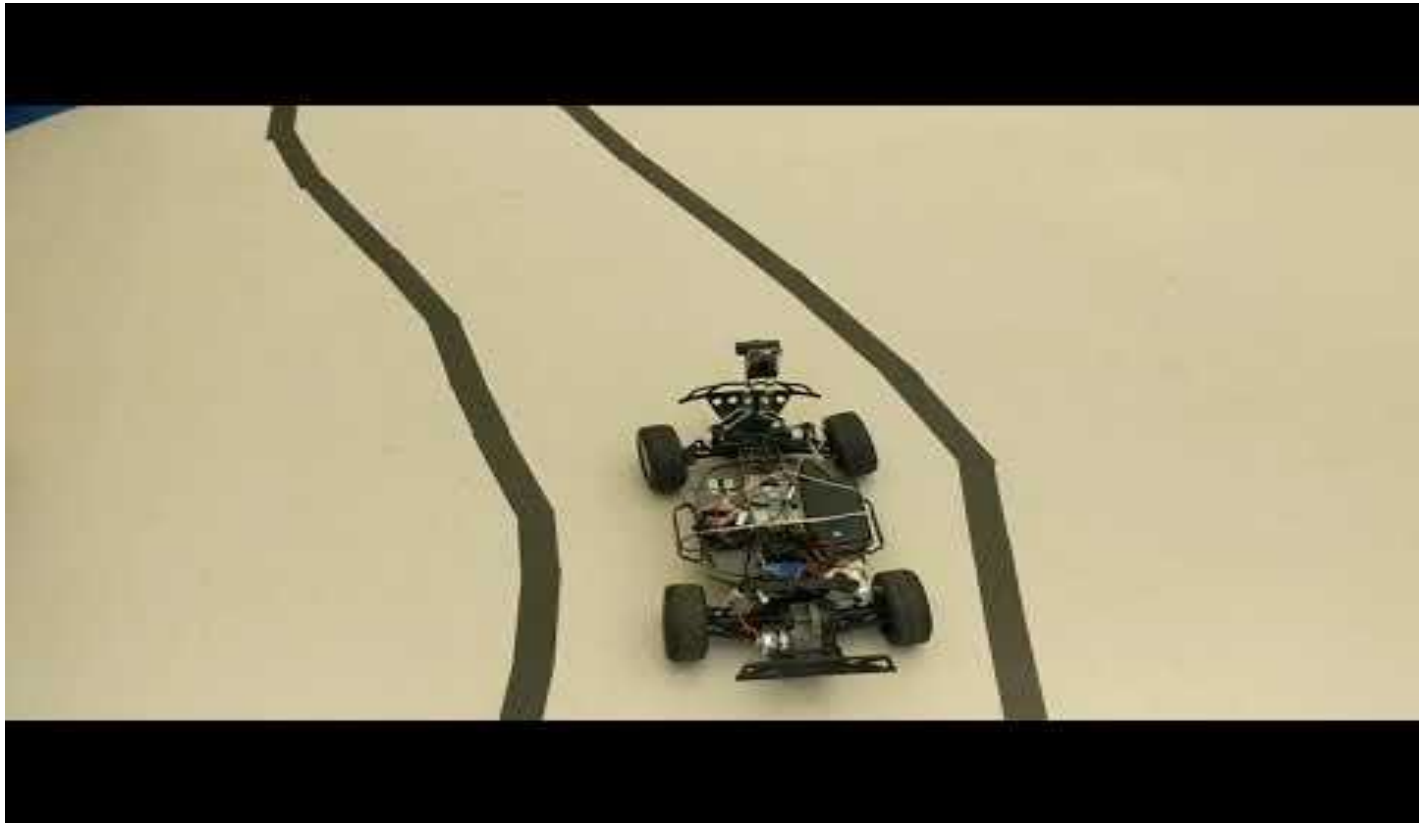
DeepNNCar – Autonomous CPS testbed



https://github.com/scope-lab-vu/deep-nn-car



Autonomous steering of DeepNNCar based on forward looking camera images



Suceptibility of AI Components

Tesla's Autopilot crash[1]

Multi-label autonomous driving dataset images

Labels
- Day
- Clear
- More Traffic
- No pedestrian

Labels
- Day
- Clear
- No Traffic
- No pedestrian

- AI components are sensitive and affected if the test image has a generative factor distribution shift from the training images. – **Out-of-Distribution Images.**
- For safety of CPS it is critical to detect OOD images and isolate the specific generative factor causing it.

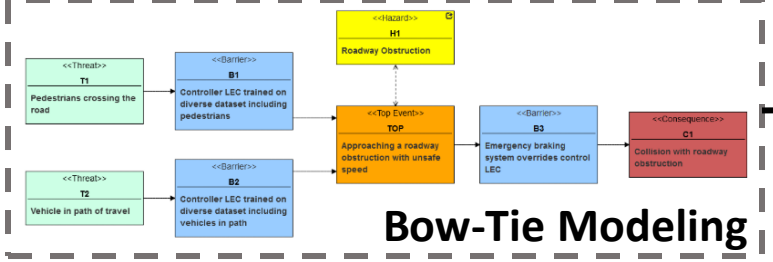
1. https://enrg.io/new-details-fatal-tesla-crash-emerge/

System Risk Assessment

System Analysis
- Design (software architecture)
- Safety Requirements
- Failure Modes & Effect Analysis

Hazards Analysis
- Identify Hazards, Threats, Barriers, Consequences

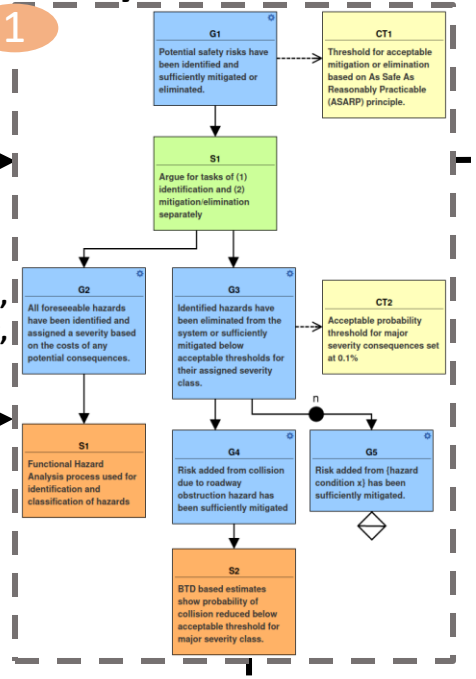
Hazard propagation



Bow-Tie Modeling

Safety Goals, Assurance claims, strategy, context, Hazards.

Safety Assurance Case (1)



BTD with conditional probabilities.

Static risk

Dynamic Assurance Case Evaluator

(4) **Decision Manager**
- e.g., controller selection

Dynamic risk estimates

(2) **ReSonAte Risk Computation**

Observed State

Sensor Observations
- Environment state
- System state

Failure Mode Hypotheses

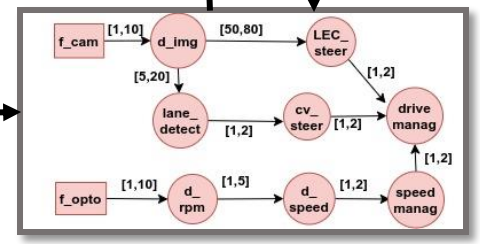
Observed State

monitor types and requirements.

Runtime Monitoring
- Assurance monitors
- Anomaly detectors

(3)

Monitor Alarms



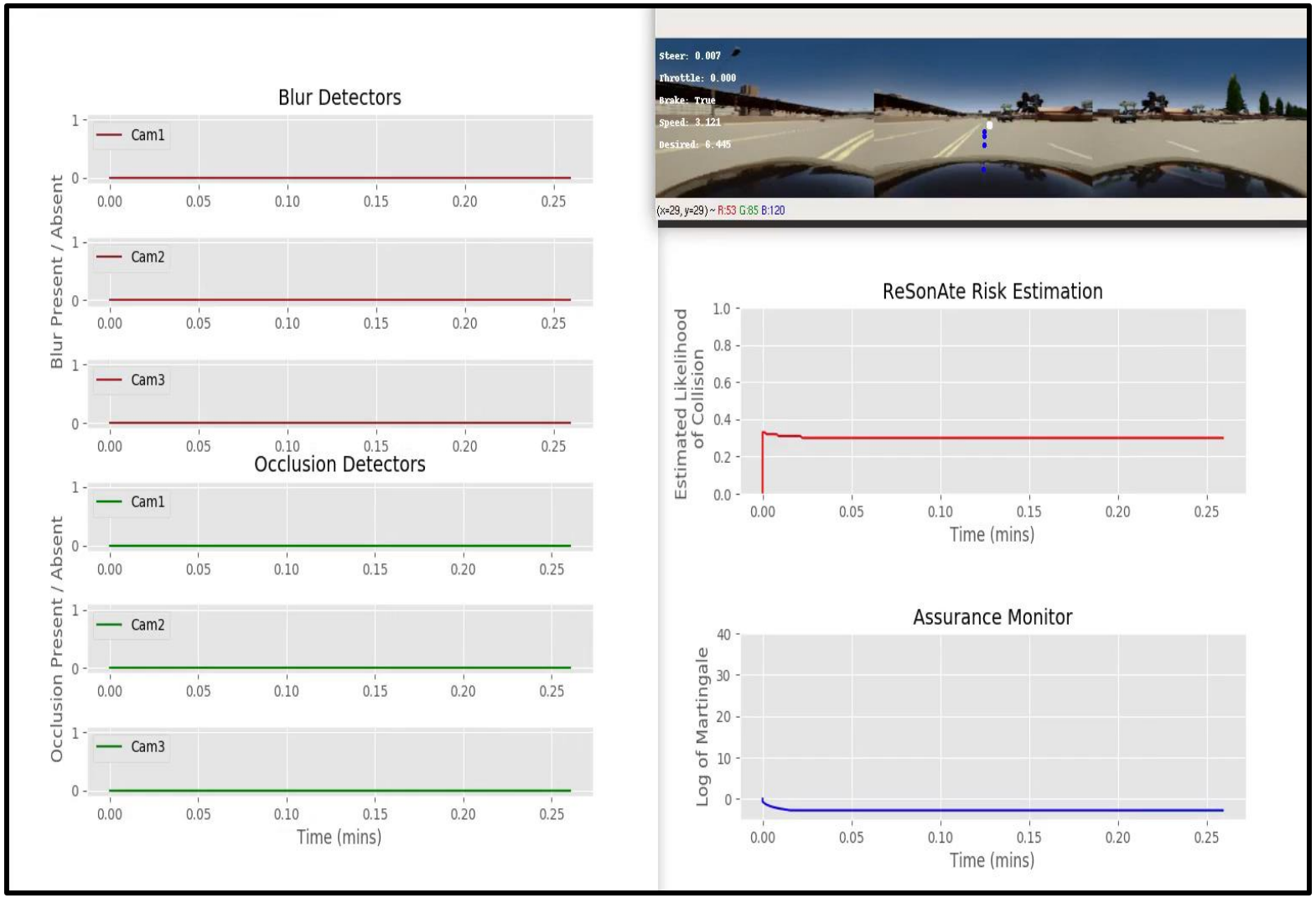
Fault Diagnosis (e.g., TFPG)

Design time steps | Runtime steps | Runtime information | Design time information

System Risk Assessment – CARLA AV

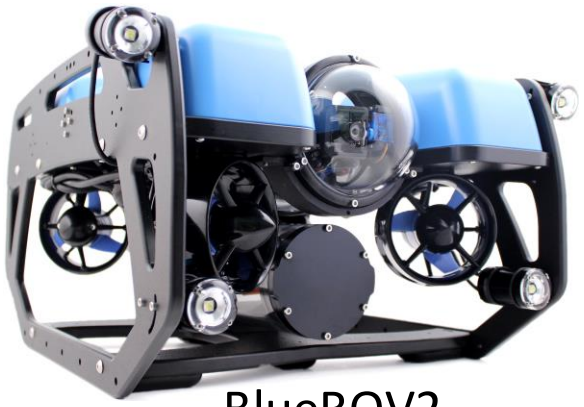
CARLA Simulator[1]

Example: Operating an AV under varying weather and sensor faults (e.g., camera faults).



1. https://leaderboard.carla.org/

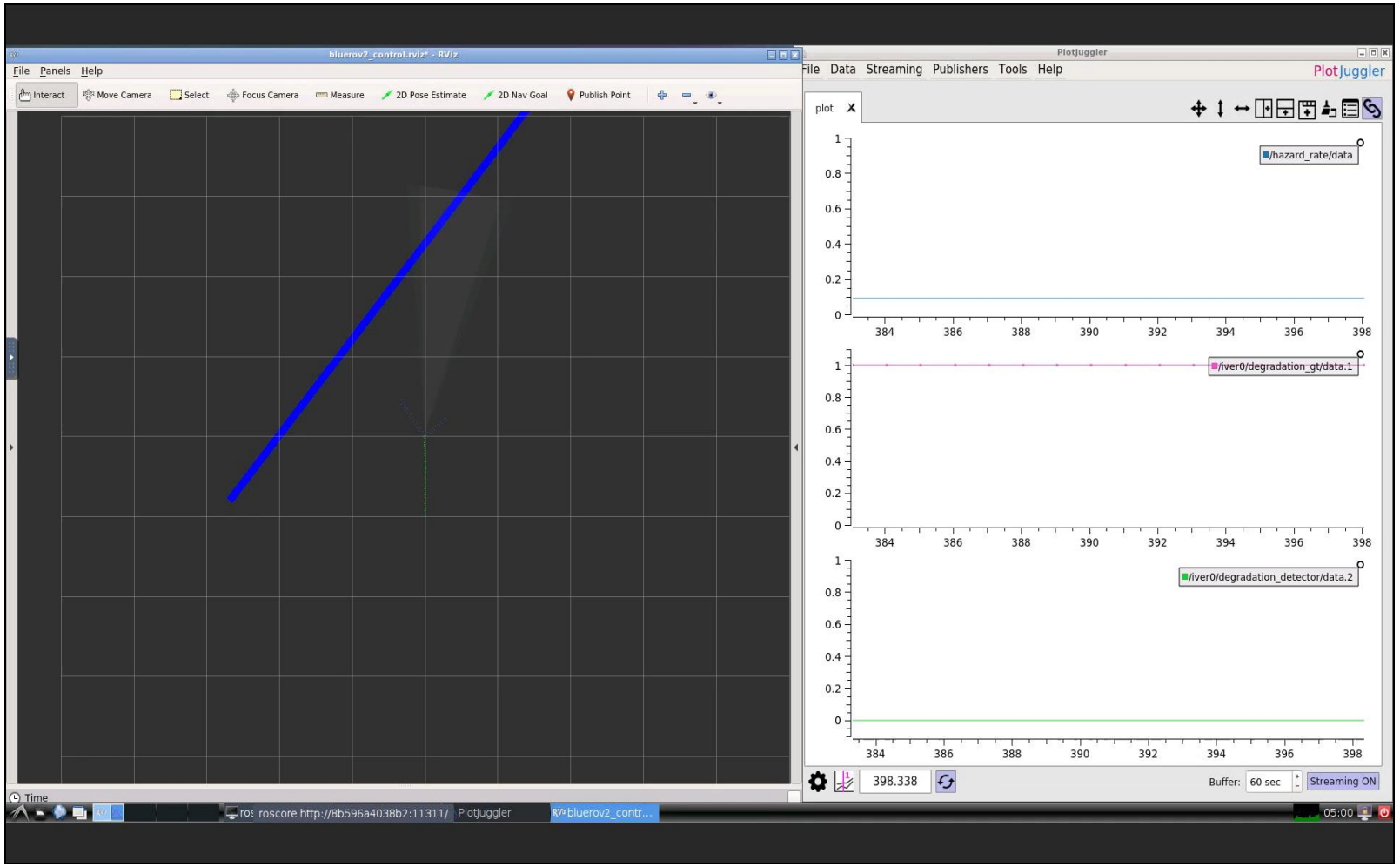
System Risk Assessment – Under Water Vehicle



BlueROV2

Example: Operating UUV in Degraded Condition.

- Perform pipe tracking + obstacle avoidance with thruster degradation.



https://bluerobotics.com/store/rov/bluerov2/

Thank You

Any Questions ?