Inherent Sensor Redundancy for Automotive Anomaly Detection

Tianjia He, Lin Zhang, Fanxin Kong, Asif Salekin

Department of EECS, Syracuse University

{the107, lzhan120, fkong03, asalekin}@syr.edu

Cyber-Physical Systems Lab

Faculty

Fanxin Kong

Interests:
Security, real-time, and energy-efficiency aspects for Cyber-Physical Systems and Internet of Things

Email: fkong03@syr.edu

Testbeds

Students

Tianjia He

Interest:
- anomaly detection
- machine learning
- cyber-physical systems
- human-computer interaction

Francis Enoch Akowuah

Interest:
- mobile cybersecurity
- cyber-physical systems security

Lin Zhang

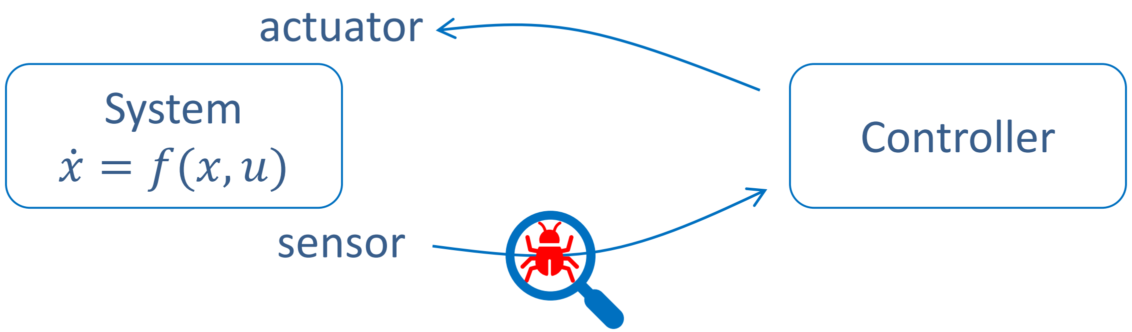
Interest:
- cyber physical systems
- checkpointing
- recovery
- control systems

Mengyu Liu

Interest:
- cyber-physical systems
- machine learning
- anomaly detection

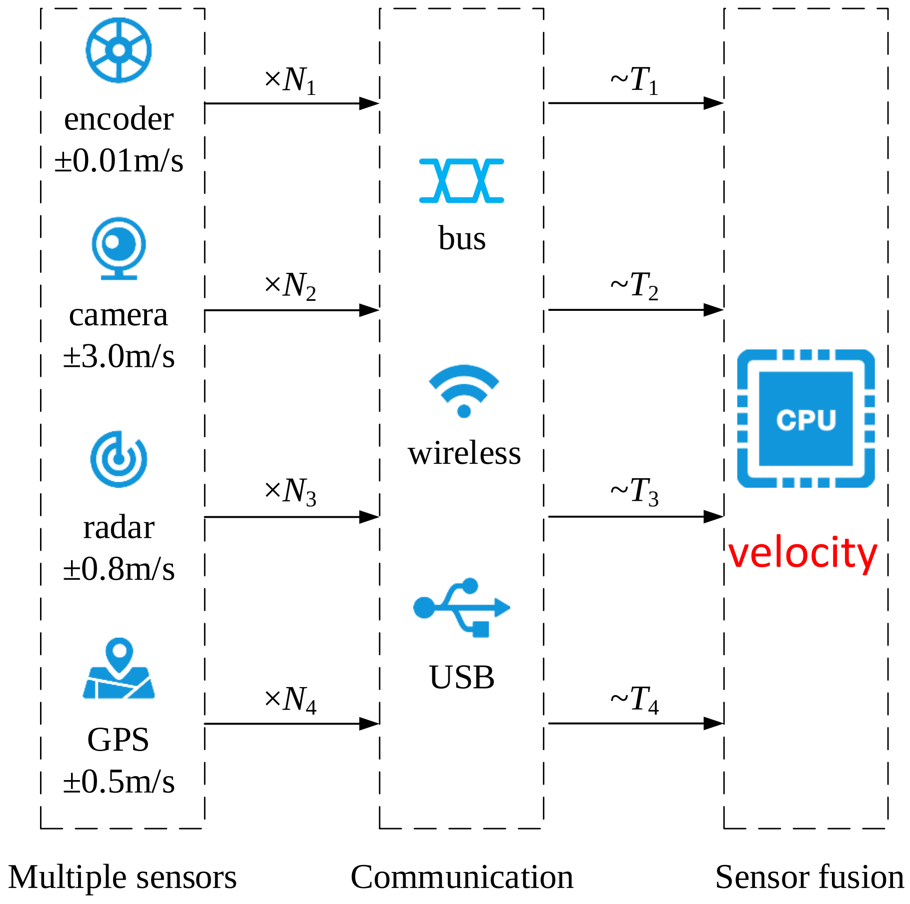
Motivation

- Security vulnerabilities in automobiles
  - increasing autonomy and connectivity
  - non-invasively compromise sensors and spoof the controller
  - exacerbated consequences on safety



- Validating sensor data before the controller acts on them
  - model-based validation
  - inherent sensor redundancy (√)

Inherent Sensor Redundancy



Multiple sensors simultaneously respond to the same physical aspect in a **related** manner.

Challenges:

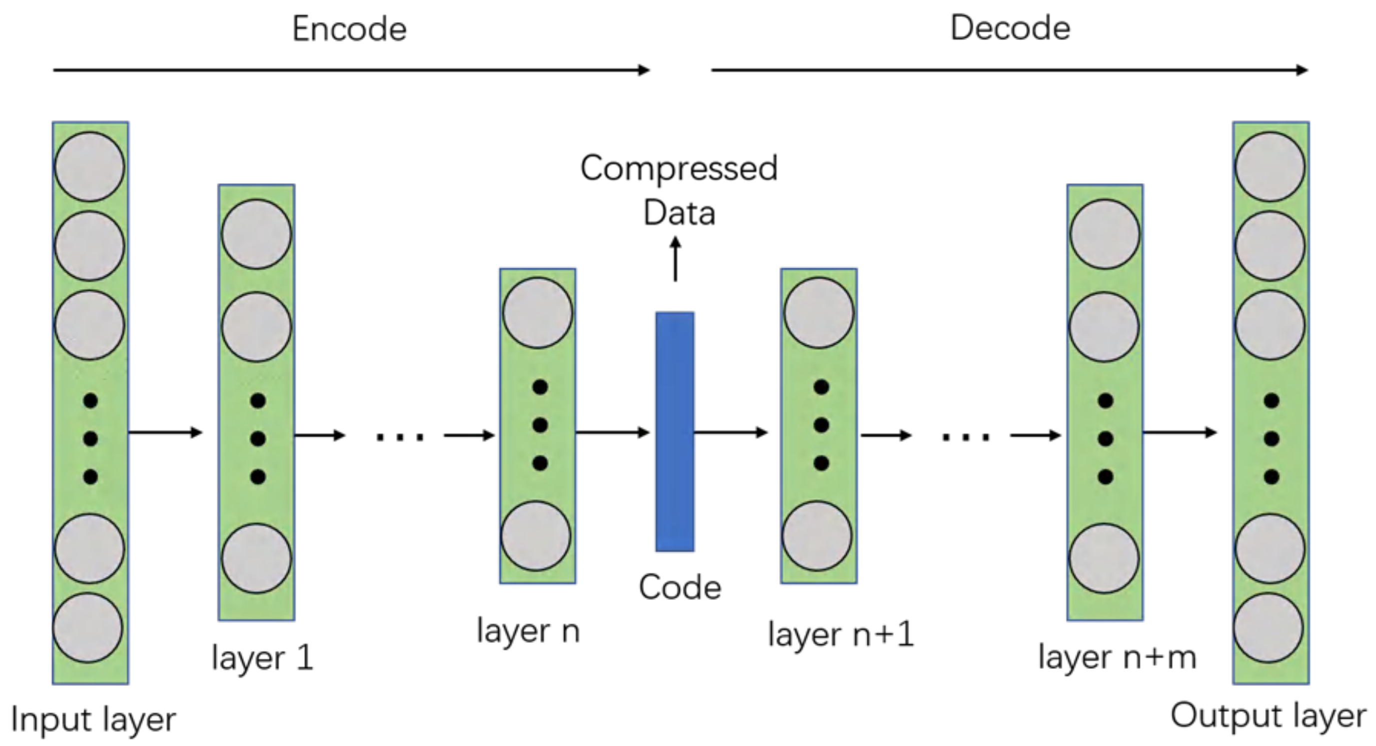
- lack of anomalous sensor data
- difficult to find a closed-form expression of the sensor relationship
- conventional assumption of disturbances on sensing

Approach

- Objective: exploits inherent redundancy among heterogeneous sensors for detecting anomalous sensor measurements
- Deep Autoencoder
  - consists of two main parts: the encoder and the decoder
  - learns a consistent pattern from vehicle sensor data in normal states
  - utilizes it as the nominal behavior for the detection
- Overview
  - Training Encoder and Decoder (using normal data)
  - Reconstruction Error Measurements
  - Threshold Estimation

Deep Autoencoder Training

Input Set of sensors data (normal)



Output set of reconstructed sensors data

$$C = \sigma_L(W^L \cdots \sigma_2(W^2\sigma_1(W^1X+b^1)+b^2)+\cdots+b^L)$$

$$\hat{X} = \sigma_M(W^M \cdots \sigma_2(W^2\sigma_1(W^1C+b^1)+b^2)+\cdots+b^M)$$

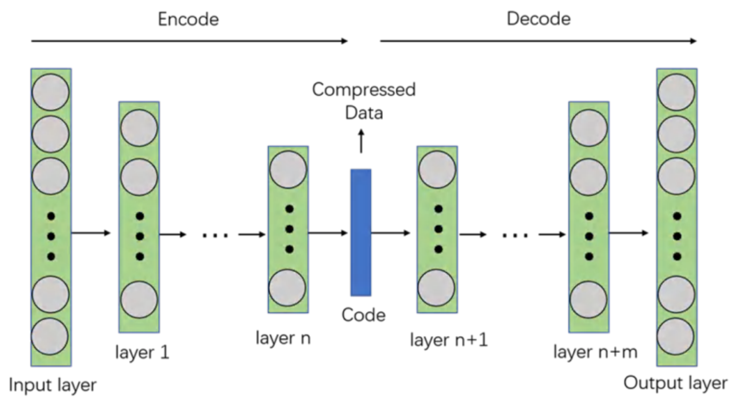
Training target : Minimize the difference between Input and Output.
Such difference are also called **Reconstruction Error**. (training loss)

Reconstruction Error Measurements

- Different Reconstruction Error
  - Mean Squared Error $D_{MSE}(X,\hat{X}) = \frac{1}{d}\sum_{i=1}^{d}(x_i - \hat{x}_i)^2$
  - Mean Square Logarithmic Error $D_{MSLE}(X,\hat{X}) = \frac{1}{d}\sum_{i=1}^{d}(log(x_i + 1) - log(\hat{x}_i + 1))^2$
  - Mean Absolute Error $D_{MAE}(X,\hat{X}) = \frac{1}{d}\sum_{i=1}^{d}|x_i - \hat{x}_i|$

Normal data → Small Reconstruction Error

Anomaly data → **Large** Reconstruction Error



Threshold Estimation

- The reconstruction error
  - Within a range for normal data
  - Define a threshold as the upper bound of the range
  - **Beyond the threshold → anomalies**
- The definition of threshold T

$$S = \frac{\sum_{i=1}^{n} D_i}{n}$$

$$T = S + 2\sqrt{\frac{\sum_{i=1}^{n} (D_i - S)^2}{n}}$$

- A relatively small and stable range can provide a meaningful threshold and be sensitive to anomalous behaviors

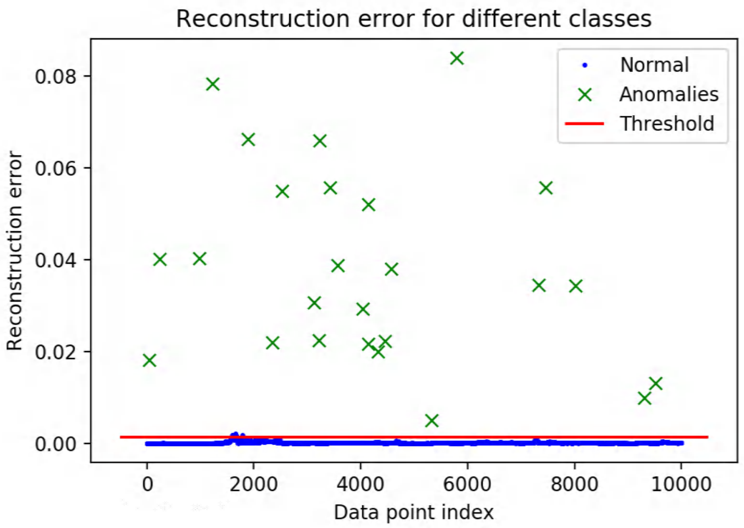
Dataset

- AEGIS dataset (real world)
  - Sensors on CAN bus
  - GPS Sensors
  - IMU Sensors
- Correlated with each other
  - acceleration pedal
  - engine RPM
  - GPS-derived speed
  - vehicle speed

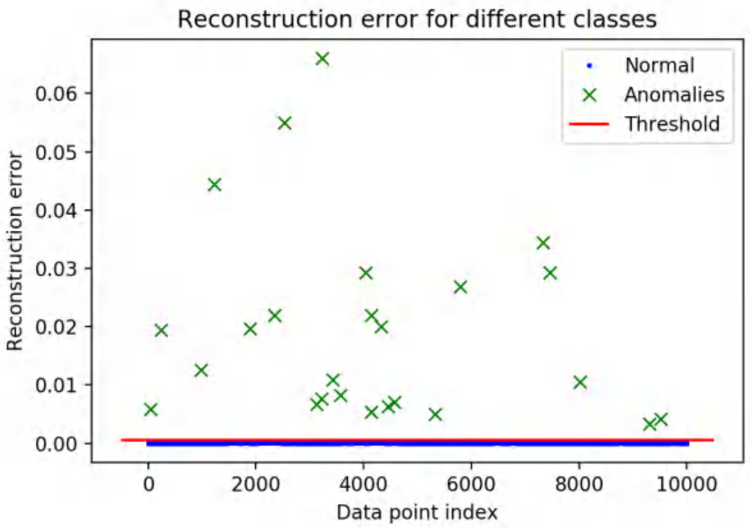
TABLE I
SENSOR DATA CONSIDERED IN THIS PAPER. SOME ABBREVIATIONS: ASR = ACCELERATION SLIP REGULATION, ACC = ACCELERATION, BRK = BRAKE, MFS = MISFIRING SYSTEM, TRQ = TORQUE, FL = FRONT LEFT, FR = FRONT RIGHT, RL = REAR LEFT, RR = REAR RIGHT, G = GRAVITY.

Sensors on CAN bus	GPS Sensors
ASR	Acceleration
AccPedal	Current sec
AirIntakeTemperature	Direction
AmbientTemperature	Distance
BoostPressure	GPS fix quality
BrkVoltage	Velocity
EngineSpeed_CAN	**IMU Sensors**
EngineTemperature	Accelerometer_X
Kickdown	Accelerometer_Y
MFS_Tip_Down	Accelerometer_Z
MFS_Tip_Up	Body_acceleration_X
SteerAngle	Body_acceleration_Y
Trq_FrictionLoss	Body_acceleration_Z
Trq_Indicated	G_force
VehicleSpeed	Magnetometer_X
WheelSpeed_FL	Magnetometer_Y
WheelSpeed_FR	Magnetometer_Z
WheelSpeed_RL	Velocity_X
WheelSpeed_RR	Velocity_Y
Yawrate	Velocity_Z

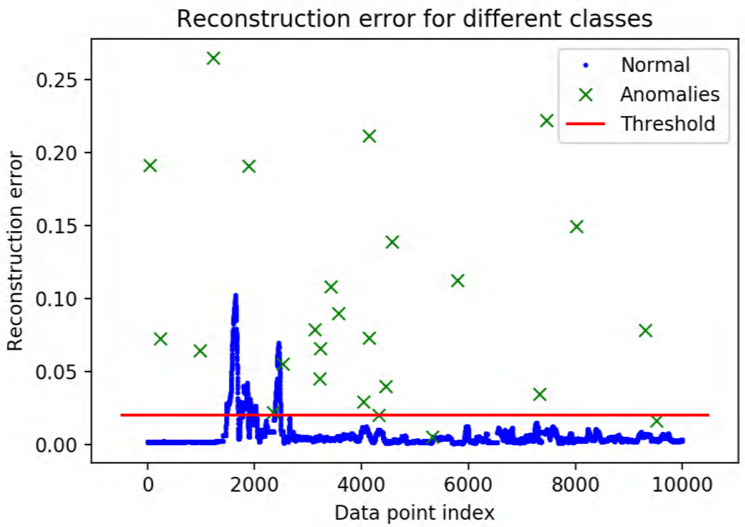
Experiment Results



MSE



MSLE



MAE

Autoencoder network: 4-layers encoder and a 4-layers decoder

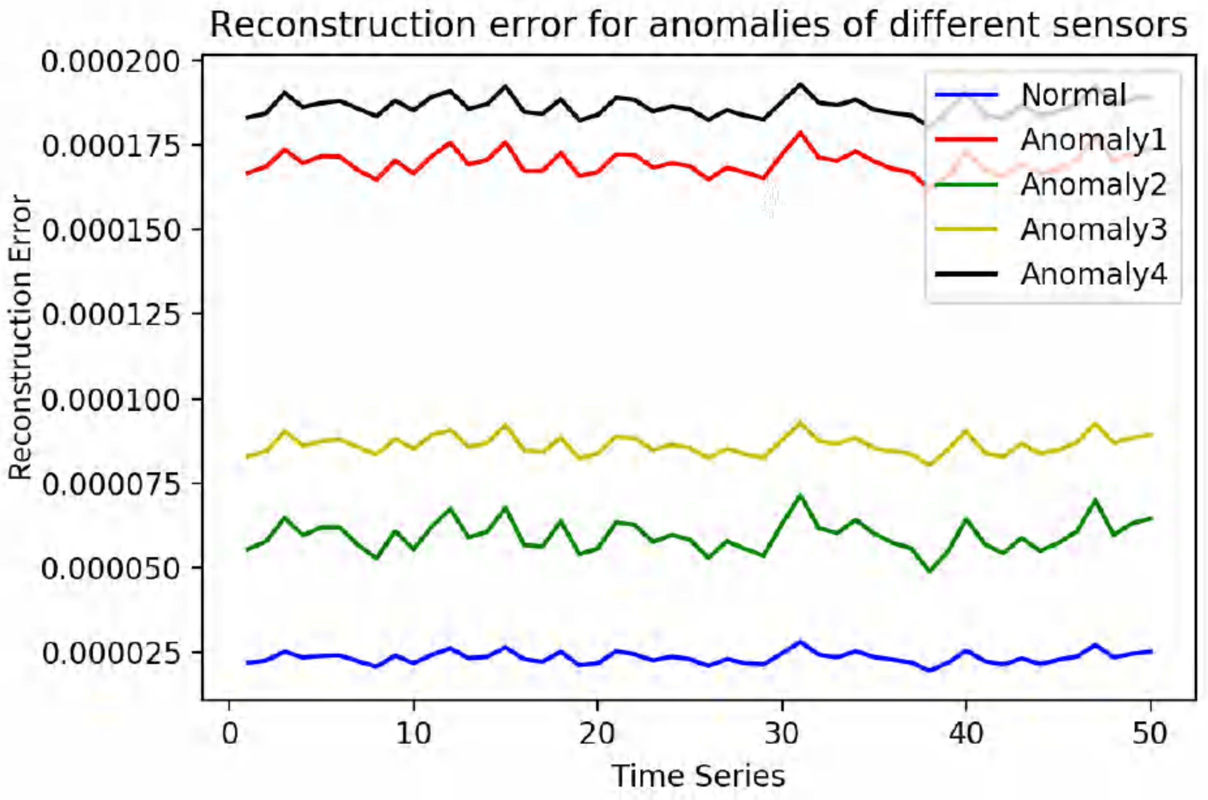
input/output size = 40

Training data : 10,000 entries of normal driving data

randomly replace 25 entries with anomalous data

Test data: 10,000 entries in 500 seconds, 25 anomalous data

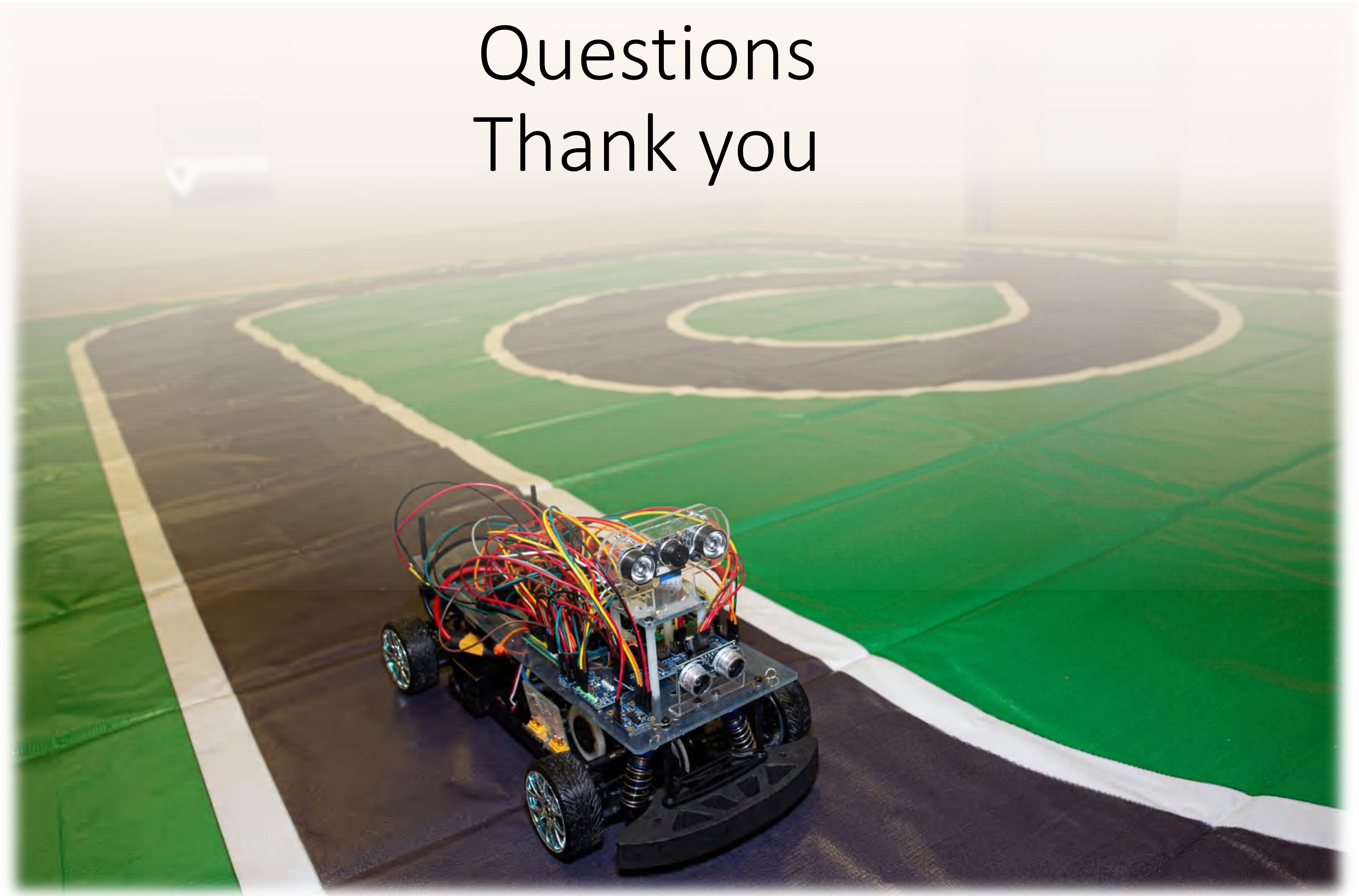
Experiment Results



Detection sensitivity to anomalous data of different sensors

- Distribution of reconstruction error of MSLE based Deep Autoencoderwith various testing samples. Each sample has a continuous anomalous data injection on a different sensor.
- Anomaly 1 - vehicle speed data
- Anomaly 2 - accPedal data
- Anomaly 3 - vehicle acceleration data
- Anomaly 4 - engine speed data

Questions
Thank you



Contact Us

Fanxin Kong

Email: fkong03@syr.edu