Reimaging Transportation Hubs with Safe and Versatile Assistant Robots

DATA COLLECTION

What data will you collect or create?

This project aims to establish an evaluation system for autonomous vehicles and connected autonomous vehicles by integrating scenario generation to generate critical scenarios and physical robots to execute the scenarios. Most data will be generated by scenario generation algorithms trained on public datasets and collected by our autonomous vehicle and robots in a closed proving ground. We will use the sensing system that is equipped on our autonomous vehicles and robots to capture real-world data, such as LiDAR point cloud and vision information. Those data will be used to boost the multi-fidelity evaluation. For example, the LiDAR data will be used to generate realistic augmented reality scenarios. The team will provide a safe and automatic evaluation program for the safety of vulnerable road users and deploy the system in the real world.

How will the data be collected or created?

We will use several scenario generation algorithms that are trained on public datasets, such as the Waymo Open dataset, nuScenes, and nuPlan, to generate critical scenarios that pose significant risks to fine-grained categories of vulnerable road users (VRUs). We will use rule-based methods to extend the scenarios from coarse-grained categories to fine-grain categories of VRUs. We will also collect data by using our autonomous vehicles and robots.

DOCUMENTATION AND METADATA

What documentation and metadata will accompany the data?

Each generated scenario instance will be accompanied by (1) geometric information, for instance, the location of each object, the road map, (2) semantic information, such as the road user type, and the attribute of each road user, (3) the goal of the evaluation, for example, LLM generated description about based on which regulation, this scenario is necessary.

ETHICS AND LEGAL COMPLIANCE

How will you manage any ethical issues?

The ethical issues will be reported to the PI's university and subject to the university's Ethics policy.

STORAGE AND BACKUP

How will the data be stored and backed up during the research?

All generated and collected data will reside on PCs and workstations that belong to the PIs' university. The data will be shared with PennSTART to deploy the physical evaluation system.
All data will be regularly backed up on external hard drives, to ensure full data recovery in the event of equipment failure. Additionally, approaches to re-create the same data will be documented with all the details, in case of catastrophic failures.

How will you manage access and security?

Data derived from this project will be retained for at least one year. The scenario dataset we created and the evaluation system we developed will be shared with PennSTART for deployment. The selected research results will be released and shared with the research community through technical reports or publications. The data will not contain any private or confidential information. The research results belong to PI's university. Selected result data and visualization will be accessible upon request to the PI and research assistants.

DATA SHARING

How will you share the data?

To the collaborator PennSTART, all the data will be shared.
With the scientific community, the PI will communicate any significant findings through publications, national and international conference presentations, and seminars, in a timely manner and throughout the project. The reported results will be available to the research community when possible, permitted, and upon request.