Project Title: Connected Vision for Increased Pedestrian Safety (CVIPS)

Recipient/Grant (Contract) Number: Carnegie Mellon University, Grant #: 69A3552344811

Center Name: Safety21 National University Transportation Center for Promoting Safety

Research Priority: Promoting Safety

Principal Investigator(s): Vijayakumar Bhagavatula Carnegie Mellon University

PI Contact: kumar@ece.cmu.edu

Project Partners:

- General Motors
- Qualcomm
- City of Pittsburgh Dept of Mobility and Infrastructure

Research Project Funding: $290,497.00

Project Start and End Date: 07-01-2023 to 06-30-2024

Project Description:

Nearly a quarter of the 1.35 million traffic fatalities in the world involve pedestrians. In the US, the number of pedestrian deaths increased from 4,280 in 2010 to an estimated 7508 in 2019, a 75% increase. Roughly one-quarter of traffic fatalities and about one-half of all traffic injuries in the US are caused at intersections where complex interactions between vehicles, pedestrians, motorcycles, and cyclists occur. This proposal is aimed at increasing the safety of vulnerable road users (VRUs) such as pedestrians, cyclists, and scooter riders, specifically at intersections. One approach to increase traffic safety is to bring more autonomy into vehicles with the goal of avoiding human-specific problems such as distractions and drunken driving. Towards this goal, computer vision algorithms are applied on data from vision sensors such as RGB cameras and LIDAR for detecting other objects in the scene. Most of these algorithms perform well on detecting larger objects such as vehicles but not as well on detecting pedestrians and other VRUs. Another challenge is that, unlike human drivers with whom pedestrians can communicate with a plethora of verbal and non-verbal cues, interacting with autonomous vehicles is a nascent concept, Further, the performance of deep learning-based vision techniques can degrade significantly in challenging conditions such as low-light environments, sun glare and severe occlusions, e.g., a pedestrian darting out from between two parked cars. In line with US DOT's vision for connected and autonomous vehicles, this proposal is aimed at developing and evaluating computer vision solutions that offer increased VRU safety by using C-V2X capabilities. Our proposed Connected Vision for Increased Pedestrian Safety (CVIPS) system is illustrated schematically in Fig. 1 (see supplement). Each agent extracts pedestrian location and trajectory information using a video vision transformer and communicates that information to other participating agents using a unified representation such as the Bird's Eye View (BEV). An example of the benefits of CVIPS is shown schematically in Fig. 2 (in the supplement) which shows a scenario with 4 vehicle cameras and 1 infrastructure (e.g., traffic light) camera. Here a pedestrian is not visible to two of the vehicle cameras but is visible to another vehicle camera and the infrastructure camera. CVIPS relies on C-V2X connectivity, and it is important to understand the impact of the C-V2X parameters (e.g., bandwidth, latency, etc.) and limitations on the data available to the multiple agents. For example, is there enough bandwidth to share the full video frames or should we share only the bounding boxes of detected VRUs? Also, achieving a unified BEV representation requires knowledge of the locations of participating cameras and the impact of location errors (e.g., caused by GNSS) needs to be studied. It is also important to consider challenging imaging conditions such as rain/snow, sun glare, nighttime, etc. Deep learning solutions can be highly demanding in storage and computational complexity and so one of our goals will be to develop light-weight implementations. Also, there are no known datasets that provide videos from multiple cameras covering the diverse range of pedestrian scenarios we propose to investigate. We propose to generate synthetic data using the high-fidelity CARLA simulator. Initial algorithm development and evaluation will be based on synthetic data, but best-performing algorithms will be tested on real data that will be collected after identifying relevant driving scenarios and obtaining the necessary IRB approvals. CVIPS project consists of the following major research tasks: (1) Creating synthetic image sequences for the VRU scenario, (2) acquiring real data after IRB approval, (3) development and testing of baseline deep learning algorithms for pedestrian detection and pedestrian trajectory estimation, (4) investigation of the impact of C-V2X parameters on the accuracy of pedestrian detection and trajectory estimation, (5) evaluating the developed algorithms under challenging imaging conditions, and (6) quantifying the increased pedestrian safety through CVIPS.

Outputs:

The anticipated outputs from this project are as follows:

- Synthetic image and image sequence datasets that includes vehicles, pedestrians and other VRUs at intersections equipped with connected cameras.

- Publications describing the deep learning algorithms for pedestrian detection and trajectory estimation in images and image sequences collected by multiple cameras at intersections, subject to C-V2X and V2X limitations.

- Software implementations of the deep learning algorithms.

- Technical report quantifying the tradeoffs between multi-agent communication parameters and pedestrian and VRU detection and trajectory estimation accuracy.

Outcomes/Impacts:

The following are the expected outcomes:

1. Creation of synthetic image and image sequence datasets (using CARLA) that includes vehicles, pedestrians and other VRUs at intersections equipped with connected cameras. These images will reflect normal operating conditions as well as challenging imaging conditions.
After appropriate vetting, this dataset will be made publicly available spurring further research into the important topic of increasing pedestrian and VRU safety at urban intersections.

2. Deep learning based pedestrian detection and trajectory estimation algorithm that offer increased robustness to challenging conditions such as occlusions, rain/snow and sun glare. The developed machine learning algorithms will have the potential to increase pedestrian and VRU safety at urban intersections.

3. Quantification of the tradeoffs between multi-agent communication parameters (e.g., communication delays, data dropouts, etc.) and pedestrian and VRU detection and trajectory estimation accuracy. This quantification will help in the understanding how the parameters of CV2-X systems can impact the VRU safety improvements achievable through collaborative vision approaches.

These outcomes will bring closer the vision of connected vehicles leading to increased safety, particularly for vulnerable road users.