


Carnegie Mellon University

Penn
University of Pennsylvania

The Ohio State University

Morgan State University

CCAC Community College of Allegheny County

The University of Texas Rio Grande Valley

# Safety Assurance and Demonstration of Connected Autonomous Vehicles

Ragunathan (Raj) Rajkumar: 0000-0001-5803-9619
Nishad Sahu: 0000-0002-0536-4538
Shounak Sural: 0009-0007-6537-8716

*Final Report – July 31, 2026*

*DISCLAIMER*

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, under [grant number 69A3552344811 / 69A3552348316] from the U.S. Department of Transportation's University Transportation Centers Program. The U.S. Government assumes no liability for the contents or use thereof.

**Technical Report Documentation Page**

| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| **4. Title and Subtitle**<br>Safety Assurance and Demonstration of Connected Autonomous Vehicles | | **5. Report Date**<br>July 31, 2026<br>**6. Performing Organization Code** |
| **7. Author(s)**<br>Ragunathan (Raj) Rajkumar: 0000-0001-5803-9619<br>Nishad Sahu: 0000-0002-0536-4538<br>Shounak Sural: 0009-0007-6537-8716 | | **8. Performing Organization Report No.** |
| **9. Performing Organization Name and Address**<br>Safety21 USDOT University Transportation Center<br>Carnegie Mellon University<br>5000 Forbes Ave.<br>Pittsburgh, PA 15213 | | **10. Work Unit No.**<br>**11. Contract or Grant No.**<br>Federal Grant # 69A3552344811 / 69A3552348316 |
| **12. Sponsoring Agency Name and Address**<br>USDOT University Transportation Center Program | | **13. Type of Report and Period Covered**<br>Final Report (July 1, 2025 – June 30, 2026)<br>**14. Sponsoring Agency Code**<br>USDOT |
| **15. Supplementary Notes** | | |
| **16. Abstract** | | |
| **17. Key Words**<br>workzone safety; autonomous vehicle safety | **18. Distribution Statement**<br>No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161. | |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of Pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | 7 | |

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized.

**Resulting Publications**


[1] N. Sahu, C. Qian, G. Cai, and S. Sural, "*WZPlanner*: Safe End-to-End Path Planning for Autonomous Driving in Work Zones," *arXiv preprint arXiv:2609.19393*, 2026. URL: https://arxiv.org/pdf/2609.19393.pdf

[2] G. T. Su, N. Sahu, S. Sural, S. Brennan, and R. R. Rajkumar, "On V2X communications for autonomous navigation of work zones and road incidents," in *2026 IEEE Intelligent Vehicles Symposium (IV)*, 2026, pp. 1226–1233. URL: https://ieeexplore.ieee.org/document/11623986

[3] X. Huang, Z. Hua, H. Zhou, S. Sural, and R. Rajkumar, "*RT-VLA*: Real-Time Vision-Language-Action Models via Knowledge Distillation," *arXiv preprint arXiv:2606.14010*, 2026. URL: https://arxiv.org/pdf/2606.14010.pdf

[4] N. Sahu, K. Panda, C. Yu, C. Qian, S. Sural, and R. Rajkumar, "*Safe2Drive*: Evaluating safe driving behaviors of E2E autonomous driving models," in *CVPR Workshops*, 2026. URL: https://openaccess.thecvf.com/content/CVPR2026W/papers/Sahu_Safe2Drive_CVPRW_2026_paper.pdf

[5] S. Sural and R. Rajkumar, "DriveSafer: End-to-End autonomous driving with safety guidance," in *CVPR Workshops*, 2026. URL: https://openaccess.thecvf.com/content/CVPR2026W/papers/Sural_DriveSafer_CVPRW_2026_paper.pdf

[6] S. Sural and R. Rajkumar, "*BEVMAPMATCH*: Multimodal BEV neural map matching for robust re-localization of autonomous vehicles," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2026. URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10126001

[7] S. Sural, "Context-Aware Autonomous Driving in Adverse Operating Conditions," Ph.D. dissertation, Carnegie Mellon University, 2026. URL: https://kilthub.cmu.edu/articles/thesis/Context-Aware_Autonomous_Driving/2026/files/pdf

[8] S. Sural, N. Sahu, and R. Rajkumar, "*WorkZone3D*: A multimodal dataset for 3D work zone perception in autonomous driving," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2026. URL: https://openaccess.thecvf.com/content/WACV2026/papers/Sural_WorkZone3D_WACV_2026_paper.pdf


**Project Description**

**WzPlanner**: Work zones are common safety-critical driving environments where lane geometry is frequently altered by temporary traffic control artifacts. Partial or complete

lane or road closures in work zones may not be reflected in onboard maps creating both perception and planning challenges for autonomous vehicles (AVs). Existing perception systems and modular AV stacks often fail to generalize well in such settings, due to structural complexities of possible road configurations and in part due to the limited availability of publicly released data with structured geometric supervision. This effort includes the **WorkZonePlan** dataset, a work zone dataset comprising 149K+ synthetic and 5K+ real-world multimodal samples with 3D annotations for lane boundaries, work zone boundaries, and driving trajectory planning options. The dataset also provides 76 closed-loop work-zone scenarios in CARLA, each replayed under three weather conditions to form 228 Bench2Drive-format evaluation routes. We also introduce WAVE (Work-zone-focused AV data generation in Virtual and rEal Environments), a semi-automated pipeline used to create the WorkZonePlan dataset. Finally, we propose BoundaryFormer (BF), a transformer-based architecture that jointly predicts lane and work zone boundary polynomials and driving trajectories.

**V2X Communications:** Autonomous vehicles (AVs) are a promising technology to reduce automotive fatalities and crashes dramatically. However, AV deployment to date has been restricted to limited operational design domains, largely because of the wide-range complexity of driving. For instance, AVs will encounter work zones while driving, because work zones are necessary to maintain, repair, and upgrade roadways. Work zones and road incidents such as crash sites are complex and safety-critical, raising significant challenges for AV perception and behaviors. Bad weather, poor lighting, weak cloud connectivity, and degraded road conditions aggravate matters further. We study how V2X communications can assist AVs to make driving safer for road users. V2X communications provide reliable and detailed information that can augment perception. We introduce a range of realistic yet complex driving environments, discuss how V2X communications improve AV safety, and capture deployment requirements for V2X. We propose three new V2X message protocols called *Lane Blockage Alert (LBA)*, *Basic Contour Message (BCM)*, and *Moving Obstacle Alert (MOA)* to address some of these requirements. We capture our experiences and results from a set of experiments conducted in test-track and real-world work zones to demonstrate the benefits of AVs using V2X in work zones.

**RT-VLA: (Real-Time Vision-Language-Action Models):** Vision-Language-Action (VLA) models have shown strong potential for end-to-end autonomous driving by jointly modeling visual perception, language reasoning, explainability and action prediction. However, their large vision language backbones and reasoning modules introduce substantial inference latency and thereby prevent their deployment in the unforgiving reality of the road networks. We propose RT-VLA, a lightweight, distilled VLA model that transfers the driving and reasoning capabilities of the state-of-the-art SimLingo model into a compact student

through multi-level supervised distillation. RT-VLA preserves language-based reasoning and supports post-hoc explanation through offline language analysis of safety-critical driving moments without adding latency to real-time control. Compared to the SimLingo teacher, RT-VLA maintains competitive closed-loop driving and language reasoning performance while reducing inference time by 44.8× in vision-only mode and 7.9× in vision+language mode. These results suggest that supervised distillation is a practical approach for building real-time, explainable VLA-style autonomous driving models.

**Safe2Drive:** Recent end-to-end (E2E) autonomous driving policies achieve high driving scores in closed-loop simulations. Yet it remains unclear whether these policies handle common safety-critical scenarios. **Safe2Drive** (S2D) is a set of Bench2Drive-aligned scenario extensions focused on three frequent families of road hazards: work zones, pedestrian jaywalking, and occluded vulnerable road users (VRUs). Safe2Drive adds 100 common but challenging scenarios and introduces **SafeDriving Score** (SDS), a safety-centric metric that augments prior evaluators with pre-crash braking, work zone-object contact, lane centering, and smoothness checks. Evaluating two state-of-the-art policies (LEAD and SimLingo) on S2D, we find that their driving scores drop sharply relative to their reported Bench2Drive baselines (LEAD: from 94.70DS on Bench2Drive to 39.95DS on S2D; SimLingo: from 85.07DS on Bench2Drive to 41.00DS on S2D) and that SDS on S2D is low (11.85 for LEAD and 15.27 for SimLingo). These results are consistent with brittle safe-driving behaviors such as poor work-zone understanding, red-light violations, and late or absent braking for pedestrians. This study highlights a lack of safe behavioral reasoning in E2E models even when tested on CARLA towns that are part of the training set. We plan to release the code and videos for all 100 S2D scenarios.

**DriveSafer**: End-to-End (E2E) autonomous driving models have shown growing capability in recent years, with performance improving on increasingly challenging benchmarks. However, modern generative E2E planners still suffer from a substantial number of catastrophic failures in safety-critical scenarios. We find that many such failures arise from violations of physical constraints and safety requirements, leading to unsafe behavior. Motivated by this finding, we focus on improving safety outcomes in generative end-to-end driving with a targeted reduction of catastrophic planning failures, instead of enhancing average planning quality. Towards this end, we propose **DriveSafer**, a failure-aware safety framework for end-to-end planners. DriveSafer explicitly steers generative planners towards safe behaviors leveraging both training-time safety constraints and inference-time safety guidance. Compared to the state-of-the-art DiffusionDrive model, on the NAVSIM benchmark, DriveSafer reduces the number of catastrophic failures (PDMS=0) by 48%, with over 65% reduction in drivable-area compliance failures.

**BEVMapMatch**: Localization in GNSS-denied and GNSS-degraded environments is a challenge for the safe widespread deployment of autonomous vehicles. Such GNSS-challenged environments require alternative methods for robust localization. In this work, we propose **BEVMapMatch**, a framework for robust vehicle re-localization on a known map without the need for GNSS priors. BEVMapMatch uses a context-aware lidar+camera fusion method to generate multimodal Bird's Eye View (BEV) segmentations around the ego vehicle in both good and adverse weather conditions. Leveraging a search mechanism based on cross-attention, the generated BEV segmentation maps are then used for the retrieval of candidate map patches for map-matching purposes. Finally, BEVMapMatch uses the top retrieved candidate for finer alignment against the generated BEV segmentation, achieving accurate global localization without the need for GNSS. Multiple frames of generated BEV segmentation further improve localization accuracy. Extensive evaluations show that BEVMapMatch outperforms existing methods for re-localization in GNSS-denied and adverse environments, with a Recall@1m of 39.8%, being nearly twice as much as the best performing re-localization baseline.

**WorkZone3D**: Work zones are essential to maintain, repair and upgrade our roadways. However, they introduce complex, dynamic and challenging environments for autonomous vehicles to navigate safely. To help address this challenge, we introduce the first publicly available, large-scale, multimodal work zone dataset collected with an autonomous vehicle consisting of multiple synchronized lidars and high-resolution cameras. Our WorkZone3D dataset covers various work zone elements such as cones, barrels, and channelizers, and provides 3D annotation boxes for these objects. We also propose a auto-annotation pipeline that can produce high-quality 3D labels, assimilating data across *Equal Contribution frames, even for rare classes which do not have pre-trained 3D object detection models to start with. We evaluate unimodal and multimodal models on our dataset, showing the critical role of sensor fusion in accurate 3D localization of such small objects at a distance, often having very few lidar points on them. Our results demonstrate the usefulness of WorkZone3D for generalization to real-world scenarios. Our code and dataset are available at

**Code and Datasets**

*WZPlanner:* *https://github.com/Nishad-Sahu/WZPlanner*

*BEVMapMatch*: https://github.com/ssuralcmu/BEVMapMatch

*Workzone 3D*: https://github.com/ssuralcmu/WorkZone3D

**ORC IDs**

1. Ragunathan (Raj) Rajkumar: 0000-0001-5803-9619

2. Nishad Sahu: 0000-0002-0536-4538

3. Shounak Sural: 0009-0007-6537-8716