DRISI

CALTRANS DIVISION OF RESEARCH, INNOVATION AND SYSTEM INFORMATION

TRANSFORMING IDEAS INTO SOLUTIONS

Research Results

Automated Vehicle (AV) Testing Data Sharing and Data Analytics

Establish an effective strategic collaboration between AV companies and Caltrans regarding data sharing.

Advanced Research

MAY 2026

Project Title: Connected and Automated Vehicle (CAV) Application Development

Task Number: 4152

Start Date: June 30, 2023

Completion Date: April 30, 2026

Task Manager:

Jaskaren Virk
Transportation Engineer (Electrical)
jaskaren.virk@dot.ca.gov

DRISI Division Chief:

Prakash Sah, PE
prakash.sah@dot.ca.gov

DRISI provides solutions and knowledge that improves California's transportation system.

WHAT WAS THE NEED?

As Connected and Automated Vehicles (CAV) continue to advance from research towards deployment, many optimists predict that fully automated vehicles will be introduced to public roadways soon. However, there are several open questions and issues that require research, planning, and resolution at State and local transportation agencies to enable the successful broad deployment.

One of the questions that needs exploration is "what transportation infrastructure improvements or modifications are needed to improve automated vehicle (AV) performance?" In order to explore this question, a survey was conducted in 2020, in which 20 AV companies were surveyed regarding infrastructure needs for large-scale AV deployment. The surveyed companies include automated car start-up companies, automated truck start-up companies, AV technology provider start-up companies, traditional automotive car manufacturers, and traditional automotive parts manufacturers.

During communications with these AV companies, many of them expressed the willingness to have interactive engagement with the public agencies and share their AV testing data for the purpose of better maintaining the public roadway infrastructures. Therefore, under the project, we have the opportunity to optimize California Department of Transportation (Caltrans) planning, operations, and maintenance by using the big data collected from AV sensors, for example, inspecting the roadway conditions and identifying the maintenance priorities.

Under this project, meaningful communication will be conducted between government agencies and the AV industry. Therefore, one goal of this project is to establish an effective strategic and collaboration between AV companies

ADA Notice: Users with accessibility issues may contact the California Department of Transportation, Division of Research, Innovation and System Information. For TTY assistance, call the California Relay Service at 711, email: pm2.communications@dot.ca.gov or write Caltrans, DRISI – MS-83, P.O. Box 942873 Sacramento, CA 94273-0001

and Caltrans regarding data sharing. The other goal is to establish the methodology and best practice of sharing AV testing data and demonstrate the value of data exchange that can be used for Caltrans' planning, operations, and maintenance.

WHAT WAS OUR GOAL?

The goal of this Agreement is to produce a report that provides specific recommendations for Caltrans regarding engagement with AV industry companies for data sharing, best practices for data sharing, and valuable applications of utilizing AV testing data for Caltrans' planning, operation, and maintenance.

WHAT DID WE DO?

This task will cover the following activities.

1. Conduct a review of the background, current state of the art, and near-term use of AV testing data generated by automated driving systems (ADS).
2. Explore and build innovative partnerships between AV companies and Caltrans regarding AV.
3. Obtain AV testing data from the companies that agree to share data for this project.
4. Conduct data transfers, data quality control, and data storage management.
5. Conduct data analytics and data visualization.

WHAT WAS THE OUTCOME?

This project demonstrates the feasibility of leveraging automated vehicle sensing data and computer vision techniques to support roadway infrastructure condition monitoring. In this project, we developed a vision-based framework for detecting road surface damage and degraded pavement markings, with a particular focus on potholes, degraded pavement signs, and degraded lane markings.

A major outcome of the project is the construction of a large-scale annotated roadway damage dataset in the US. Through manual filtering, mask-based annotation, and an iterative semi-automated annotation pipeline, the project generated more than 20,000 annotated images for model training and validation. The use of mask-based polygon annotations allowed the model to capture irregular road damage boundaries more accurately than conventional bounding-box labels, which is especially important for pothole and pavement marking deterioration analysis.

The trained YOLOv12m-seg model achieved strong detection and segmentation accuracy across multiple roadway damage classes, including potholes, degraded pavement signs, degraded lane markings, and manholes. The overall model performance reached 0.849 mAP50 for bounding-box detection and 0.808 mAP50 for mask segmentation, confirming that YOLO-based segmentation is suitable for scalable roadway damage detection. The model also maintained reliable performance under challenging lighting and weather conditions, especially when image enhancement preprocessing was applied.

The trained model also showed promising robustness under different visual conditions, including normal lighting, low-light scenes, shadows, and rainy conditions. The prediction examples indicate that the model can identify roadway damage targets even in challenging environments. Additional preprocessing using image enhancement, particularly CLAHE-based contrast enhancement, further improved detection confidence and size estimation stability under low-light and rainy conditions.

Building upon the detection results, this project proposed a stereo vision-based pothole size estimation method. The method integrates instance segmentation, depth sensing, camera geometry, and calibration-based correction to estimate pothole width, length, and area. Experimental results showed that raw perspective-based estimation can produce large errors, particularly for pothole length estimation. After applying a distance-dependent

The contents of this document reflect the views of the authors, who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the California Department of Transportation, the State of California, or the Federal Highway Administration. This document does not constitute a standard, specification, or regulation. No part of this publication should be construed as an endorsement for a commercial product, manufacturer, contractor, or consultant. Any trade names or photos of commercial products appearing in this document are for clarity only.

© Copyright 2026 California Department of Transportation. All rights reserved.

correction factor, the estimation accuracy improved substantially. In representative test cases, the corrected width and length errors were generally reduced to below 5–10%, and many cases achieved errors within approximately 5% compared with ground-truth measurements.

The robustness tests further demonstrated that image enhancement can significantly improve the reliability of pothole size estimation in adverse conditions. Without enhancement, confidence and size estimation accuracy degraded notably as measurement distance increased, especially in rainy or low-light conditions. With enhancement, detection confidence remained more stable, and length estimation errors were consistently reduced. These findings suggest that the proposed enhancement and calibration pipeline can support field deployment.

The framework can be integrated into AV company data pipelines to process newly collected road condition videos, automatically detect pavement defects, estimate pothole dimensions, and generate roadway damage assessment outputs for Caltrans maintenance planning.

WHAT IS THE BENEFIT?

Results of this project will help Caltrans in planning for optimizing Caltrans planning, operations, and maintenance efforts by using the big data collected from AV sensors.

The contents of this document reflect the views of the authors, who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the California Department of Transportation, the State of California, or the Federal Highway Administration. This document does not constitute a standard, specification, or regulation. No part of this publication should be construed as an endorsement for a commercial product, manufacturer, contractor, or consultant. Any trade names or photos of commercial products appearing in this document are for clarity only.

© Copyright 2026 California Department of Transportation. All rights reserved.