SJSU SAN JOSÉ STATE UNIVERSITY

CALIFORNIA STATE UNIVERSITY Chico

Development of a Framework for Identifying Asphalt Pavement Cracking Distresses Using Machine Learning

DingXin Cheng, PhD, PE

Project 2435
March 2026

Introduction

Cracking is one of the most common and damaging distresses in asphalt pavements. Cracking reduces structural integrity, ride quality, and safety while increasing maintenance costs of the roadways we all depend upon. Traditional visual inspection methods are labor-intensive, dangerous, subjective, inefficient, and sometimes create significant data gaps for road management agencies. This research developed a cost effective, automated framework for detecting asphalt pavement cracking using machine learning and computer vision. The study explored the use of deep learning architecture, particularly the YOLO (You Only Look Once) model, to accurately identify and classify pavement distresses such as alligator, block, and longitudinal cracks from high-resolution imagery. The primary goal was to enhance the accuracy, speed, and consistency of pavement condition assessments and provide a scalable solution that could be deployed through vehicle-mounted or drone-based imaging systems. By giving agencies better data on pavement conditions, the system enables more informed decisions about when and where to repair roads, saving time and money.

Study Methods

The research team developed and tested a prototype machine learning (ML) framework to automate the identification of asphalt pavement cracking. The framework combines deep learning and image processing, trained using a large dataset of annotated pavement images collected from California roadways. Each image was labeled according to standardized distress types—based on ASTM, FHWA, Caltrans, and MTC manuals—to ensure consistency across different severity levels and distress categories. A convolutional neural network (VGG16) was used for image classification—essentially teaching the system to recognize what type of crack appears in each image—while a YOLO model performed real-time object detection and localization, identifying where cracks appear on the pavement and outlining them in the image. The models were trained and validated using a 70/20/10 data split, with data augmentation techniques applied to improve robustness under varying lighting and texture conditions, so the system can still perform well even when images are taken at different times of day or on different pavement surfaces.

Model performance was evaluated using precision, recall, and mean average precision (mAP) metrics, which are standard measures that indicate how often the system correctly detects cracks and how accurately it labels them. The best performing model achieved a recall of 0.73, precision of 0.70, and an mAP50 score of approximately 0.71, demonstrating a reliable detection of multiple crack types under complex field conditions. The research also incorporated visualization outputs called bounding boxes and overlays to verify detection accuracy and support quantitative assessment. This integrated ML approach reduces the need for manual inspection and allows for rapid, objective, and repeatable pavement evaluations suitable for large-scale deployment.

Findings

The developed framework successfully demonstrated that deep learning-based methods can accurately identify and classify common asphalt pavement cracking distresses. In other words, the system can "look" at road images and recognize different cracking types. The trained models—especially the YOLO detector—achieved high accuracy in detecting alligator, block, and longitudinal/transverse cracks, even under challenging conditions such as variable lighting and surface shadows. Performance evaluation showed that the model maintained over 70% mean average precision (mAP) across major crack types and successfully localized cracks on pixel-level maps, meaning it could pinpoint exactly where cracks occurred on the pavement surface. Compared to traditional manual surveys, the automated system provided substantial improvements in safety, efficiency, consistency, and scalability.

It processed high resolution pavement images rapidly, making it compatible with both vehicle-mounted and drone-based data collection platforms. Additionally, the model's ability to integrate distress type, severity, and location data enables more objective assessments of pavement condition and supports quantitative mapping within pavement management systems, which enables agencies to better decide when and where to repair roads. However, some limitations were observed, such as missed detections of minor or low contrast cracks. These can be addressed through expanded datasets and further optimization of model parameters. Overall, the findings confirm that integrating machine learning into pavement assessment provides a practical pathway toward cost-effective, data-driven maintenance management to support safe roads.

AI- and machine-learning models achieved high detection accuracy for major asphalt pavement crack types, enabling rapid, objective, and scalable pavement assessments for safer roadways.

Practice Recommendations

Transportation agencies should adopt AI- and ML-based methods to complement or replace manual pavement inspections. Automated distress detection can substantially reduce inspection time, improve accuracy, and enable consistent statewide pavement assessments. The framework developed in this research can be integrated into existing Pavement Management Systems (PMS) to automate data input, optimize maintenance schedules, and prioritize rehabilitation efforts. Future implementation should focus on expanding the detection scope to include additional distresses—such as potholes, rutting, and raveling—and refining models for mobile and edge-based applications. Policymakers and engineers should support the development of standardized digital distress databases and invest in training for data interpretation and system integration. These practices will enhance decision-making, lower maintenance costs, and extend pavement service life through timely, data-driven interventions.

About the Author

Dr. Ding Xin (Ding) Cheng is a Professor in the Department of Civil Engineering, California State University, Chico; Director of the California Pavement Preservation (CP2) Center; and the Director of the Tire Rubber Technology Center. He has extensive experience in HMA materials and pavement preservation on both asphalt and concrete pavements. Ding has co-managed or managed more than $10 million in research projects funded by Caltrans, the California Department of Resources Recycling and Recovery (CalRecycle), the Metropolitan Transportation Commission (MTC), and other agencies and industry.

To Learn More

For more details about the study, download the full report at transweb.sjsu.edu/research/2435

MTI is a University Transportation Center sponsored by the U.S. Department of Transportation's Office of the Assistant Secretary for Research and Technology and by Caltrans. The Institute is located within San José State University's Lucas Graduate School of Business.