Holistically Identifying Road Complexity and Relating it to Fatal Crashes

Recipient/Grant (Contract) Number: University of Massachusetts – Amherst, 69A3552348301

Center Name: New England University Transportation Center (NEUTC)

Research Priority: Promoting Safety

Principal Investigator(s): Dr. Shannon Roberts, Meng Wang

Project Partners: University of Massachusetts - Amherst

Research Project Funding: $60,000 (Federal), $60,000 (Non-Federal)

Project Start and End Date: 1/1/2024 – 2/1/2025

Project Description: This project developed and tested a practical way to measure "road complexity" and relate it to where crashes are more likely. The team combined three kinds of information from naturalistic driving: what's in the scene (e.g., vehicles, signs, buildings), how the vehicle was moving (speed/acceleration), and basic context (road type, weather, time of day). They trained a two-stage model: first, an encoder learned a compact "complexity" signal from those inputs; second, that signal, together with the original inputs, was used to predict crash density. The approach was built with real driving video and vehicle data and linked to historical crash hotspots. It showed that using the learned complexity signal improved prediction accuracy compared with using raw inputs alone.

US DOT Priorities: The work supported U.S. DOT priorities by advancing data-driven methods to reduce crashes. By turning video, vehicle telemetry, and simple context into a reliable indicator of complex roadway conditions—and tying that indicator to where crashes cluster—the project strengthened tools for proactive safety analysis, screening, and targeting of countermeasures. It also informed safer automated and driver-assist functions by identifying features of scenes associated with higher crash density.

Outputs:

- A cleaned, labeled dataset that merged scene content, vehicle movement, and context for thousands of roadway frames.
- A two-stage modeling pipeline that learned a road-complexity signal and used it to predict crash density.
- Benchmarks showing higher accuracy when the learned complexity features were included versus baselines without them.
- Feature-importance summaries that highlighted which scene and driving factors most influenced predicted crash density.
- Documentation of data preparation, modeling steps, and evaluation procedures.

Outcomes/Impacts:

- **Better prediction:** Adding the learned complexity features improved accuracy over standard models that used only raw inputs.
- **Actionable signals:** The model highlighted factors linked to higher crash density (e.g., certain road types, speed patterns, and visual scene elements), offering practical clues for site screening and countermeasure selection.
- **Scalability:** Automated scene labeling and compact complexity features reduced manual effort, making large-area analyses more feasible.
- **Technology relevance:** Results can inform driver-assist and automated systems by flagging challenging scenes and supporting safer behavior planning.
- **Path to practice:** The workflow can be adapted to other regions and datasets to prioritize locations for engineering, enforcement, or operational strategies.

Final Research Report: Final report is posted here: https://www.umass.edu/neutc/projects/holistically-identifying-road-complexity-and-relating-it-fatal-crashes

U.S. Department of Transportation

Office of the Secretary of Transportation