GSoc TRIP Test

2024

ML4Sci

The purpose of TRIP is to use machine learning to analyze, quantify, and predict driver behavior.

Although some kinematic quantifications of risky driving from naturalistic driving exist, the cost, ethical concerns, and lack of experimental controls make naturalistic driving research inefficient and limited. As a result, driving simulator studies provide more controlled assessments of human behavior and performance in a variety of roadway situations not possible to replicate or assess in real world on-road observations of driving.

The data from driving simulation scenarios do not readily group to quantify driving behavior as safe, risky, drowsy, etc. Therefore, it is essential to understand how raw driving simulator data can quantify and predict driving behavior in an effort to reduce motor vehicle crash risk.

Overview

One of the most successful methods to objectively and safely measure driving behavior is to put a human driver in an immersive, high-fidelity driving simulator and simulate a specific real-world driving scenario. The driving simulator measures a variety of kinematic-based metrics (e.g., velocity, acceleration), vehicle sensor metrics (e.g., brake pressure, throttle pressure), and roadway position metrics (e.g., lane offset) sampled at 30-60 Hz. Additionally, observation cameras record the simulated roadway environment and the driver themselves during the simulated drive. Two relevant sources of data are output: 1) A .dat file with raw simulator data; and 2) A MP4 file with video synchronized with the .dat file.

Task

Using the two sources of data, create a qualification and/or quantification of driver behavior, including how the driver reacts to relevant stimuli in the simulated driving environment. Qualify and/or quantify: 1) How risky the driver is; 2) How risky the simulated driving environment is; 3) How attentive the driver is; and 4) how drowsy the driver is. Using machine learning and the MP4 file, 5) outline or annotate all objects in the simulated driving environment that may pose a safety risk to the driver and/or warrant the driver's attention. Use as many parameters in the two sources of data you think are necessary.

https://drive.google.com/drive/folders/1WqMkmoLr4exin6uPYQxgqr-IjE-vou5E?usp=drive_link

Deliverables

- Qualifications/Quantifications to the 5 items noted in the task as an accessible link
- Script saved in the Github link you share that includes the process, model creation, training, and testing. The script should be able to be run from start to finish without user intervention. It should produce clear metrics and/or plots. Data may be augmented as you please, but that process should be automated to allow us to test on other .dat and MP4 files withheld.
- If any pre-trained models are used, add a link to a Google Drive that contains the pre-trained model(s).
- A copy of your CV/Resume as an accessible link.

Please email the above information to Dr. Andrea Underhill (atunderhill@ua.edu), using the subject line “GSoC “[specific title of GSoC project you are interested in]”. Include in the body of the email links to information as specified in items above.

Metrics

- Accuracy of annotations
- Testing on withheld data

It is understood that not everyone would have expertise in driving behavior and/or driving simulators. As a result, perfection is not expected. Creativity and problem solving related to thought and effort is the primary concern. The ability to deliver the qualifications/quantifications in a manner or terminology that would be sensible in driving research is additionally rewarded.