

Using Big Data and Machine Learning Techniques for Transportation System and Efficiency Studies

Fengxiang Qiao

Innovative Transportation Research Institute, Texas Southern University, Houston, USA

Abstract

The integration of large-scale data analytics and machine learning has fundamentally transformed modern Intelligent Transportation Systems (ITS). Over the past years, we have focused on deploying the advanced data-driven methodologies to solve complex challenges in transportation system safety, traffic efficiency, and asset management. We have applied sophisticated machine learning algorithms such as random forest, k-nearest neighbor (KNN), Convolutional Neural Network (CNN), transfer learning, and reinforcement learning to evaluate the performance and structural resilience of roadside safety devices, characterize driving behaviors, and enhance traffic operational efficiencies. A representative recent work involves intelligent infrastructure management, notably applying state-of-the-art real-time object detection models—such as YOLOv12 and CNN—for the automated condition rating of traffic signals and pavement markings, thereby streamlining inspection processes and optimizing maintenance workflows. Through utilizing massive datasets, we can proactively identify deficiencies and predict critical maintenance needs before physical failures occur and leverage big data and Vehicle-to-Infrastructure (V2I/V2X) communication models to enhance the decision-making process for transportation systems. By bridging computational methodologies with empirical transportation engineering, these studies have demonstrated how big data and machine learning techniques can effectively optimize infrastructure longevity and ultimately foster safer, highly efficient, and resilient transportation networks.