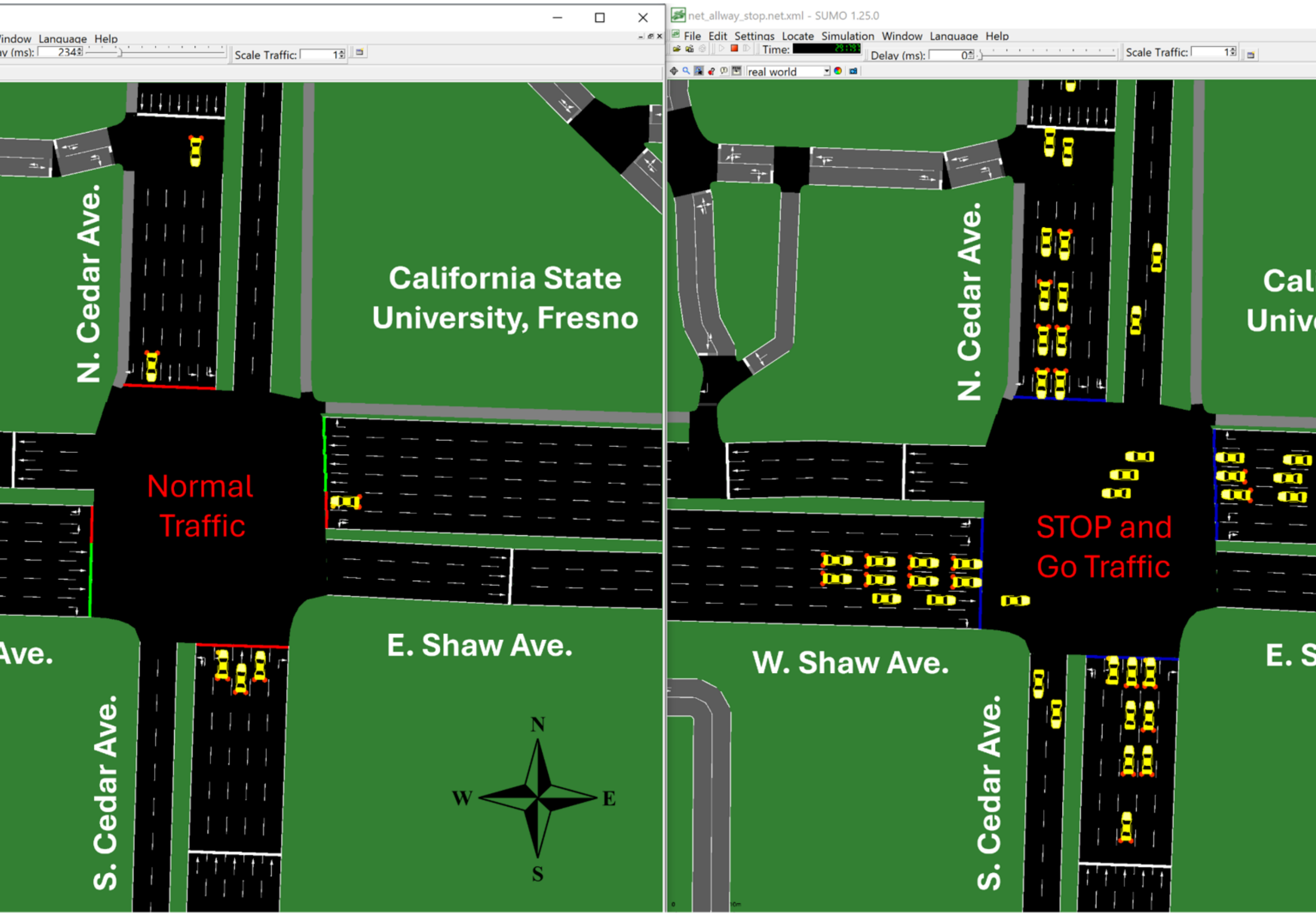
Project 2537 | July 2026

# SMART-TWIN: Systematic Modeling and Real-Time Transportation with Digital Twins

Hovannes Kulhandjian




California State University Transportation Consortium

# MINETA TRANSPORTATION INSTITUTE

Founded in 1991, the Mineta Transportation Institute (MTI), an organized research and training unit in partnership with the Lucas College and Graduate School of Business at San José State University (SJSU), increases mobility for all by improving the safety, efficiency, accessibility, and convenience of our nation's transportation system. Through research, education, workforce development, and technology transfer, we help create a connected world. MTI leads the California State University Transportation Consortium (CSUTC) funded by the State of California through Senate Bill 1 and the Climate Change and Extreme Events Training and Research (CCEETR) Program funded by the Federal Railroad Administration. MTI focuses on three primary responsibilities:

### Research

MTI conducts multi-disciplinary research focused on surface transportation that contributes to effective decision making. Research areas include: active transportation; planning and policy; security and counterterrorism; sustainable transportation and land use; transit and passenger rail; transportation engineering; transportation finance; transportation technology; and workforce and labor. MTI research publications undergo expert peer review to ensure the quality of the research.

### Education and Workforce Development

To ensure the efficient movement of people and goods, we must prepare the next generation of skilled transportation professionals who can lead a thriving, forward-thinking transportation industry for a more connected world. To help achieve this, MTI sponsors a suite of workforce development and education opportunities. The Institute supports educational programs offered by the Lucas Graduate School of Business: a Master of Science in Transportation Management, plus graduate certificates that include High-Speed and Intercity Rail Management and Transportation Security Management. These flexible programs offer live online classes so that working transportation professionals can pursue an advanced degree regardless of their location.

### Information and Technology Transfer

MTI utilizes a diverse array of dissemination methods and media to ensure research results reach those responsible for managing change. These methods include publication, seminars, workshops, websites, social media, webinars, and other technology transfer mechanisms. Additionally, MTI promotes the availability of completed research to professional organizations and works to integrate the research findings into the graduate education program. MTI's extensive collection of transportation-related publications is integrated into San José State University's world-class Martin Luther King, Jr. Library.

Report 26-23

# SMART-TWIN: Systematic Modeling and Real-Time Transportation with Digital Twins

Hovannes Kulhandjian

July 2026

A publication of the

**Mineta Transportation Institute**

Created by Congress in 1991

College of Business
San José State University
San José, CA 95192-0219

# Technical Report Documentation Page

| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| 26-23 | | |
| **4. Title and Subtitle**<br>SMART-TWIN: Systematic Modeling and Real-Time Transportation with Digital Twins | | **5. Report Date**<br>July 2026<br>**6. Performing Organization Code** |
| **7. Authors**<br>Hovannes Kulhandjian http://orcid.org/0000-0002-8983-6757 | | **8. Performing Organization Report**<br>CA-MTI-2537 |
| **9. Performing Organization Name and Address**<br>Mineta Transportation Institute<br>College of Business<br>San José State University<br>San José, CA 95192-0219 | | **10. Work Unit No.**<br>**11. Contract or Grant No.**<br>SB1-SJAUX_2023-26 |
| **12. Sponsoring Agency Name and Address**<br>State of California SB1 2017/2018<br>Trustees of the California State University<br>Sponsored Programs Administration<br>401 Golden Shore, 5th Floor<br>Long Beach, CA 90802 | | **13. Type of Report and Period Covered**<br>**14. Sponsoring Agency Code** |
| **15. Supplemental Notes**<br>10.31979/mti.2026.2537 | | |

**16. Abstract**

Urban transportation networks are increasingly affected by everyday congestion, non-recurring disruptions, and limited real-time visibility into what is happening on the road. Transportation agencies require advanced tools that support faster, proactive management and data-driven infrastructure decisions rather than passive monitoring. This report presents SMART-TWIN, an Artificial Intelligence (AI)–driven Digital Twin framework designed for real-time traffic operations, disruption detection, and performance-based decision-making support. The SMART-TWIN architecture integrates microscopic traffic simulation (detailed traffic modeling), long-horizon synthetic demand generation simulating real travel patterns, supervised machine learning, and performance analysis within a unified and reproducible workflow. A corridor-scale Digital Twin was developed for the Shaw–Cedar intersection in Fresno, California, using OpenStreetMap data and the SUMO simulation platform. An 11-year synthetic traffic dataset (2020–2030) was generated to emulate realistic traffic patterns over time and support early-stage deployment prior to full sensor availability. Supervised machine learning models were trained to classify operational conditions, including normal traffic, signal failure–induced stop-and-go conditions, and one- and two-lane closures. A Random Forest classifier achieved very high accuracy under controlled experimental conditions. To support operational decision-making, the framework translates detected traffic states into clear, measurable impacts, including estimated congestion buildup and delay-related costs using value-of-time principles. Scenario and sensitivity analyses demonstrate the compounding effects of capacity loss and control failure on congestion growth and user delay, as well as the system's ability to detect disruptions early and reliably under moderate demand uncertainty. The results show that AI-enhanced Digital Twins can provide actionable insights into traffic conditions and operational impacts prior to full-field deployment. The proposed framework is scalable to larger corridors and networks and is designed for future integration with real-time sensor data and adaptive traffic control strategies. SMART-TWIN provides a practical foundation for next-generation Digital Twin applications that support proactive traffic management and data-driven infrastructure planning to make traveling safer and more efficient for all.

| 17. Key Words | 18. Distribution Statement | | |
|---|---|---|---|
| Digital twin, intelligent transportation systems, traffic operations, microscopic traffic simulation, sumo, machine learning, incident detection, traffic disruption analysis. | No restrictions. This document is available to the public through The National Technical Information Service, Springfield, VA 22161. | | |
| **19. Security Classif. (of this report)** | **20. Security Classif. (of this page)** | **21. No. of Pages** | **22. Price** |
| Unclassified | Unclassified | 45 | |

Form DOT F 1700.7 (8-72)

DOI: 10.31979/mti.2026.2537

Mineta Transportation Institute
College of Business
San José State University
San José, CA 95192-0219

Email: mineta-institute@sjsu.edu

transweb.sjsu.edu/research/2537

# Acknowledgments

This study was supported by both the CSU Transportation Consortium and the Fresno State Transportation Institute. Any opinions, findings, conclusions, and recommendations expressed in this material are those of the author and do not necessarily reflect the views of these institutes. The author would like to thank his student assistant, Carson Gray, an undergraduate student from the Electrical and Computer Engineering (ECE) department at California State University, Fresno, for working on this interesting research project and implementing and testing the final proposed prototype. The author would also like to thank Editing Press for editorial services, as well as MTI staff, including Executive Director, Karen Philbrick, PhD, Deputy Executive Director, Hilary Nixon, PhD, Graphic Designer, Alverina Eka Weinardy, and Executive Administrative Assistant, Jill Carter.

# Contents


| | |
|---|---|
| **Acknowledgments** | **vi** |
| **List of Figures** | **viii** |
| **Executive Summary** | **1** |
| **1. Introduction** | **3** |
| **2. Literature Review** | **5** |
| **3. SMART-TWIN Architecture** | **7** |
| **4. System Model and Implementation** | **9** |
| **5. AI-Driven Data Preparation and Analysis Pipeline** | **15** |
| **6. Scenario-Based Operational Analysis** | **19** |
| **7. Queueing-Based Backlog Proxy and Cost Proxy** | **22** |
| **8. Robustness and Sensitivity Analysis** | **27** |
| **9. Early-Warning Detection Capability** | **29** |
| **10. Role of AI-Based State Classification in SMART-TWIN** | **30** |
| **11. Conclusion** | **32** |
| **Abbreviations and Acronyms** | **33** |
| **Bibliography** | **34** |
| **About the Author** | **36** |

# List of Figures


Figure 1. SMART-TWIN Framework from Data Acquisition to AI-Driven Simulation and Decision Support 8

Figure 2. OSM-to-SUMO Pipeline for Shaw–Cedar Network Construction and Validation 10

Figure 3. OpenStreetMap (OSM) Extraction of the Shaw–Cedar intersection, defining lane Connectivity and Topology for SMART-TWIN Network Construction 11

Figure 4. Synthetic Shaw–Cedar Dataset (2020–2030) for Digital Twin Calibration and AI Training 12

Figure 5. Confusion Matrices Comparing Logistic Regression and Random Forest Performance for Traffic State Classification 17

Figure 6. AI Classification Performance Comparison Between Logistic Regression and Random Forest Models 17

Figure 7. Baseline SUMO Simulation Under Normal Operating Conditions 19

Figure 8. SUMO Simulation of a Two-Lane Closure on Shaw Avenue (Right) Compared with Normal Operation (left) 20

Figure 9. SUMO Simulation of Stop-and-Go Traffic Due to Signal Failure (Right) Compared with Normal Operation (Left) 21

Figure 10. SUMO Vehicles-in-Network a 24-hour Period Across Scenarios 23

Figure 11. Average Network Delay Proxy Across Operational Scenarios 24

Figure 12. Normalized Daily User-Cost Proxy Across Operational Scenarios 25

# Executive Summary

Urban transportation systems are facing increasing operational pressure due to growing travel demands, recurring congestion, aging infrastructure, and limited real-time situational awareness. Non-recurring events such as lane closures, signal failures, incidents, and work zones often produce disproportionate impacts on traffic delays, reliability, and user cost. Transportation agencies require tools that move beyond traditional monitoring and visualization to support proactive, data-driven operational decision-making.

This research presents SMART-TWIN, an Artificial Intelligence (AI)–driven Digital Twin framework designed to support real-time traffic operations, disruption detection, and infrastructure decision analysis. The proposed system integrates microscopic traffic simulation, synthetic demand generation, supervised machine learning, and performance-based operational metrics within a unified and reproducible architecture. The framework is demonstrated through a corridor-scale Digital Twin of the Shaw–Cedar intersection in Fresno, California.

SMART-TWIN combines multiple components to emulate real-world traffic conditions and evaluate operational scenarios. OpenStreetMap data were converted into a detailed microscopic network using the SUMO simulation platform. A long-horizon synthetic traffic dataset (2020–2030) was developed to replicate realistic daily, weekly, and seasonal demand patterns, enabling Digital Twin operation during early deployment stages before full sensor availability. Traffic conditions were simulated at 10-minute intervals, and lane-level volumes and operational features were used to train supervised machine learning models.

The system classifies four operational states: normal operation, stop-and-go conditions caused by signal failure, one-lane closure, and two-lane closure. A Random Forest classifier achieved high classification accuracy (approximately 99.9%) under controlled experimental conditions. Beyond state detection, SMART-TWIN translates operational conditions into decision-relevant performance indicators using queueing-based backlog estimation and delay-to-cost conversion based on value-of-time principles. This approach enables agencies to understand not only what disruption is occurring, but also its operational and economic impact.

Scenario-based analysis demonstrates the nonlinear effects of capacity loss and control failure on congestion growth and user delay. Sensitivity analysis shows that the system maintains reliable disruption detection under moderate demand uncertainty, while impact metrics scale appropriately with traffic intensity. Early-warning analysis indicates that operational disruptions can be identified within one to two 10-minute intervals, allowing potential proactive response before severe congestion develops.

The results demonstrate that AI-enhanced Digital Twins can function as practical decision-support tools for transportation operations. By linking traffic state recognition to interpretable performance and cost metrics, SMART-TWIN bridges the gap between simulation, analytics, and operational planning. Potential applications include proactive traffic management, infrastructure investment prioritization, incident response planning, and evaluation of work-zone or capacity-reduction scenarios.

Although this Phase I study relies on synthetic demand and controlled scenario reconstruction, the framework is designed for integration with real-time sensor data and scalable deployment across corridors and network-level systems. Future work will focus on field validation, live data integration, and the incorporation of adaptive control strategies to support real-time operational optimization.

Overall, SMART-TWIN provides a scalable and practical foundation for next-generation transportation Digital Twins that support proactive management, improved system reliability, and data-driven infrastructure decision-making.

# 1. Introduction

Urban transportation systems are experiencing increasing operational stress due to population growth, rising travel demand, aging infrastructure, and expanding expectations for safety, reliability, and sustainability. Congestion continues to impose significant economic and societal costs through travel delays, fuel consumption, emissions, and reduced productivity. While recurring congestion remains a major concern, non-recurring disruptions, such as traffic incidents, work zones, lane closures, special events, and traffic signal failures, often produce disproportionate impacts on travel time variability and network performance (FHWA, 2021). These events are difficult to anticipate and manage using traditional monitoring and control approaches.

Transportation agencies are increasingly seeking data-driven tools that enable proactive system management rather than reactive response. Advances in sensing, connected infrastructure, and data analytics have expanded the availability of traffic data; however, many operational systems remain limited to visualization dashboards or historical performance analysis. There is a growing need for integrated platforms that combine real-time data, predictive modeling, and operational evaluation to support informed decision-making under dynamic conditions.

Digital Twin technology has emerged as a promising framework for addressing these challenges. A transportation Digital Twin is a virtual representation of physical infrastructure and traffic processes that integrates data, simulation, and analytics to monitor system conditions and evaluate operational strategies (Gopalakrishnan & Peeta, 2021). Recent large-scale initiatives have demonstrated the feasibility of city-scale Digital Twin environments, such as the Virtual Singapore project (Smart Nation Singapore, 2018). However, many existing implementations emphasize visualization or offline scenario analysis and lack tight integration between microscopic simulation, machine learning–based state inference, and interpretable performance metrics that directly support operational and infrastructure decisions (Ali et al., 2023; Ge et al., 2024).

This research addresses these gaps through the development of SMART-TWIN, an Artificial Intelligence (AI)–driven Digital Twin framework for real-time traffic operations and decision support. The proposed system integrates microscopic traffic simulation, synthetic demand generation, supervised machine learning, and queueing-based performance evaluation within a unified and reproducible architecture. The framework is designed to support early-stage deployment scenarios, where full sensor coverage may not yet be available, while remaining extensible to real-time data integration and adaptive control.

A corridor-scale case study is demonstrated using the Shaw–Cedar intersection in Fresno, California to demonstrate the capabilities of the framework. Using a long-horizon synthetic dataset designed to emulate realistic traffic patterns, the system detects operational disruptions such as lane closures and signal failures and translates these conditions into interpretable measures of congestion growth, delay, and user cost. By linking traffic state detection to decision-relevant performance indicators, SMART-TWIN enables transportation practitioners to evaluate operational impacts and prioritize mitigation strategies.

The key objectives of this research are the following:

- Develop a scalable Digital Twin architecture that integrates microscopic simulation and AI-based operational state detection.
- Create a calibration-ready synthetic data pipeline to support early deployment and controlled system evaluation.
- Quantify operational impacts of disruptions using queueing-based backlog estimation and delay-to-cost metrics.
- Demonstrate the feasibility of early-warning detection and scenario-based decision support at the corridor level.

This study relies on synthetic demand and simulated environments; while this enables controlled evaluation, real-world deployment may introduce additional uncertainties such as sensor noise, driver variability, and incomplete data. These factors will be addressed in future field validation.

The remainder of this report is organized as follows. Section II reviews related work in transportation Digital Twins and AI-based traffic analysis. Section III describes the SMART-TWIN system architecture. Section IV presents the data generation and simulation environment. Section V details the machine learning and analysis methodology. Section VI presents classification performance and model evaluation results. Section VII introduces scenario-based operational analysis. Section VIII presents the queueing-based backlog and cost proxy formulation. Section IX evaluates robustness and sensitivity to demand and disruption severity. Section X discusses early-warning detection capability. Section XI describes the role of AI-based state classification within the SMART-TWIN framework, followed by conclusions.

# 2. Literature Review

This section provides context for the development of SMART-TWIN by reviewing prior work in Digital Twins, simulation-based approaches, and AI-driven traffic analysis. Digital Twin (DT) technology has gained increasing attention in transportation research as a framework for integrating data, simulation, and analytics to improve system monitoring and decision-making. A transportation Digital Twin creates a dynamic virtual representation of physical infrastructure and traffic operations, enabling agencies to evaluate system performance, test operational strategies, and anticipate the impacts of disruptions under varying conditions (Gopalakrishnan & Peeta, 2021).

## Digital Twins in Transportation Systems

Recent studies have examined the role of Digital Twins in Intelligent Transportation Systems (ITS) from both conceptual and implementation perspectives. Recent studies highlight the potential of DTs to support real-time monitoring, predictive analytics, and adaptive control by integrating technologies such as Internet of Things (IoT) sensing, cloud computing, and artificial intelligence (Ali et al., 2023; Ge et al., 2024; Xing et al., 2025). Despite this potential, many current DT implementations remain limited to visualization platforms or offline "what-if" analysis rather than operational decision support.

Several large-scale initiatives have demonstrated the feasibility of urban Digital Twins. For example, the Virtual Singapore project provides a city-scale digital environment for urban planning and infrastructure analysis (Smart Nation Singapore, 2018). While these efforts illustrate the scalability of DT concepts, they often focus on planning and visualization rather than real-time traffic operations.

## Microscopic Simulation-Based Digital Twins

Microscopic traffic simulation has become a key component of operational Digital Twin frameworks. Simulation platforms such as SUMO enable detailed modeling of vehicle interactions, lane-level behavior, and traffic control strategies (Lopez et al., 2018). Research has shown that coupling microscopic simulation with real-world traffic data can support real-time system monitoring and predictive analysis (Kusić et al., 2023).

Additional work has integrated SUMO-based Digital Twins with visualization engines such as Unity to provide interactive operator interfaces and enhanced situational awareness (Rundel et al., 2023). Other studies have explored Digital Twin–based municipal traffic control systems that iteratively evaluate and implement control actions using real-time sensing data (Pradhananga et al., 2025). These efforts demonstrate the operational value of runtime simulation but often focus primarily on control evaluation rather than disruption detection and decision-oriented performance assessment.

## AI and Machine Learning for Traffic State Analysis

Artificial intelligence and machine learning techniques have been widely applied to traffic prediction, anomaly detection, and incident identification. Data-driven models can capture nonlinear relationships among traffic variables and improve the detection of abnormal conditions compared with traditional threshold-based methods (Lv et al., 2015). In the context of Digital Twins, machine learning enables automated operational state inference, supporting early detection of congestion patterns, capacity reductions, and control failures.

However, many AI-based traffic studies focus on prediction accuracy without linking model outputs to operational decision metrics. For transportation agencies, understanding the practical impact of detected conditions, such as expected delay growth or user cost, is essential for prioritizing mitigation actions.

## Gap in Existing Research

Despite these advances, existing Digital Twin systems often lack integration between simulation, AI-based inference, and decision-oriented performance metrics. In addition, many approaches assume fully instrumented environments and do not address early-stage deployment scenarios.

## Contribution of This Research

SMART-TWIN addresses these gaps by integrating microscopic simulation, a calibration-ready synthetic data pipeline, supervised traffic-state classification, and queueing-based delay-to-cost metrics within a unified, reproducible Digital Twin framework. Unlike prior approaches, the proposed system explicitly links AI-based disruption detection to actionable operational and economic indicators, enabling practical decision support for transportation agencies.

# 3. SMART-TWIN Architecture

SMART-TWIN is designed as a modular, end-to-end Digital Twin framework that integrates data ingestion, microscopic traffic simulation, artificial intelligence–based state inference, and performance evaluation within a unified operational workflow. The architecture is intended to support both early-stage deployment using synthetic demand and future integration with real-time field data.

At a high level, SMART-TWIN operates as a closed-loop system in which traffic inputs drive a microscopic simulation environment, machine learning models infer operational conditions, and the resulting system states are translated into interpretable performance and decision-support metrics.

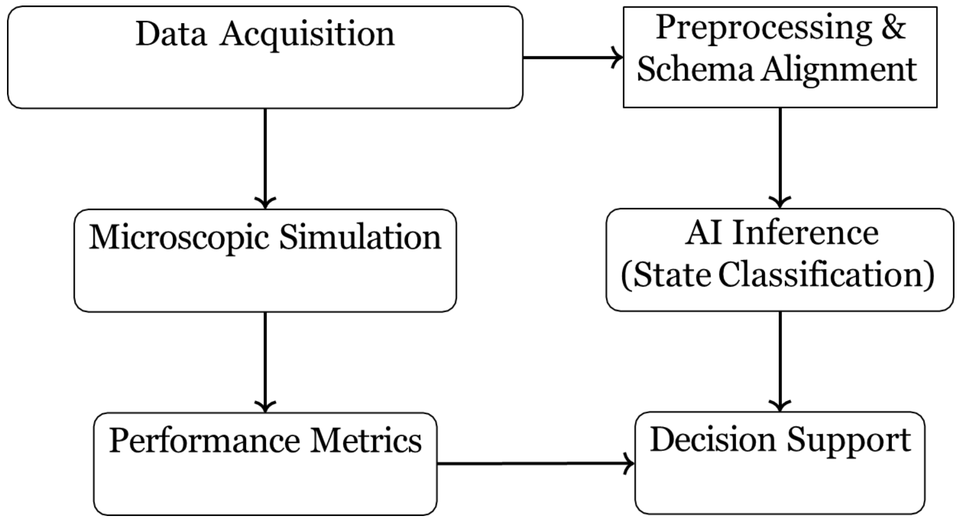
## Overall Architecture

The SMART-TWIN framework consists of four primary components:

- Data acquisition and preprocessing
- Microscopic simulation and live emulation
- AI-based operational state inference
- Performance evaluation and decision support

Figure 1 illustrates the SMART-TWIN framework from data acquisition through AI-driven simulation and decision support. These components operate sequentially but are designed to support continuous updates, enabling near–real-time Digital Twin operation.

Figure 1. SMART-TWIN Framework from Data Acquisition to AI-Driven Simulation and Decision Support



# 4. System Model and Implementation

## System Model and Digital Twin State Representation

The SMART-TWIN system is modeled in discrete time, where each time step k corresponds to a 10-minute interval. The system state is defined as:

$$\mathrm{x}_k = [\mathrm{q}_k^\top, \mathrm{n}_k^\top, \mathrm{v}_k^\top, \mathrm{d}_k^\top, \theta_k^\top]^\top. \tag{1}$$

This state representation captures both traffic conditions and contextual variables used for analysis within the Digital Twin. Here, $\mathrm{x}_k$ denotes the DT state vector, where $\mathrm{q}_k^\top$ is the vector of lane/approach queue lengths, $\mathrm{n}_k^\top$ denotes lane occupancies or vehicle counts, $\mathrm{v}_k^\top$ represents mean speeds, $\mathrm{d}_k^\top$ captures delay/travel-time statistics, and $\theta_k^\top$ are model/context variables (e.g., weather/event indicators and scenario flags).

The system evolves according to:

$$\mathrm{x}_{k+1} = f(\mathrm{x}_k, \mathrm{u}_k, \mathrm{w}_k) \tag{2}$$

where $\mathrm{u}_k$ represents inputs (demand, control actions), and $\mathrm{w}_k$ captures disturbances such as demand variability or incidents.

Observed outputs from the Digital Twin are given by:

$$\mathrm{y}_k = h(\mathrm{x}_k) + \mathrm{v}_k \tag{3}$$

where $\mathrm{v}_k$ represents measurement or modeling noise.

## AI-Based Operational State Inference

SMART-TWIN incorporates supervised machine learning to infer operational conditions from observed system outputs.

The operational state space is defined as:

$$\mathcal{S} = \{\text{Normal, Stop-and-Go, 1-lane closure, 2-lane closure}\} \tag{4}$$

The inferred state is given by:

$$\hat{s}_k = g(\boldsymbol{\phi}_k) \tag{5}$$

where $\phi_k$ represents the extracted feature vector derived from $\mathrm{y}_k$.

A Random Forest classifier is used to approximate g(·), enabling robust classification under nonlinear and heterogeneous traffic conditions.

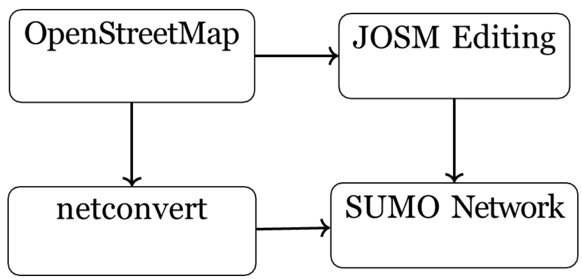
## Digital Twin Construction and SUMO Integration

The Digital Twin is implemented using the Simulation of Urban Mobility (SUMO) platform, with roadway geometry derived from OpenStreetMap (OSM) and manually refined to ensure accurate lane connectivity, turning movements, and signal control logic. Figure 2 illustrates the OSM-to-SUMO network construction pipeline.

Traffic demand is injected into the simulation using the TraCI interface, enabling real-time interaction between system inputs and simulation outputs. At each time step, the Digital Twin produces observations $\mathbf{y}_k$, which serve as a measurable representation of the system state.

This simulation-driven framework enables controlled evaluation of traffic dynamics under varying demand levels and disruption scenarios.

Figure 2. OSM-to-SUMO Pipeline for Shaw–Cedar Network Construction and Validation




## Synthetic Data Generation and Live Emulation

A Python-based demand generator was developed to emulate realistic traffic patterns over extended time horizons. The model produces temporally varying demand traces that capture the following:

- Morning and evening peak periods
- Weekday and weekend variations
- Weather-induced demand fluctuations
- Event-driven traffic surges

The generator was extended to produce a long-horizon dataset spanning 2020–2030, enabling robust system evaluation under diverse traffic conditions. Demand is generated at a 10-minute resolution to capture short-term dynamics relevant to congestion formation and disruption analysis.

Figure 3. OpenStreetMap (OSM) Extraction of the Shaw–Cedar intersection, defining lane Connectivity and Topology for SMART-TWIN Network Construction

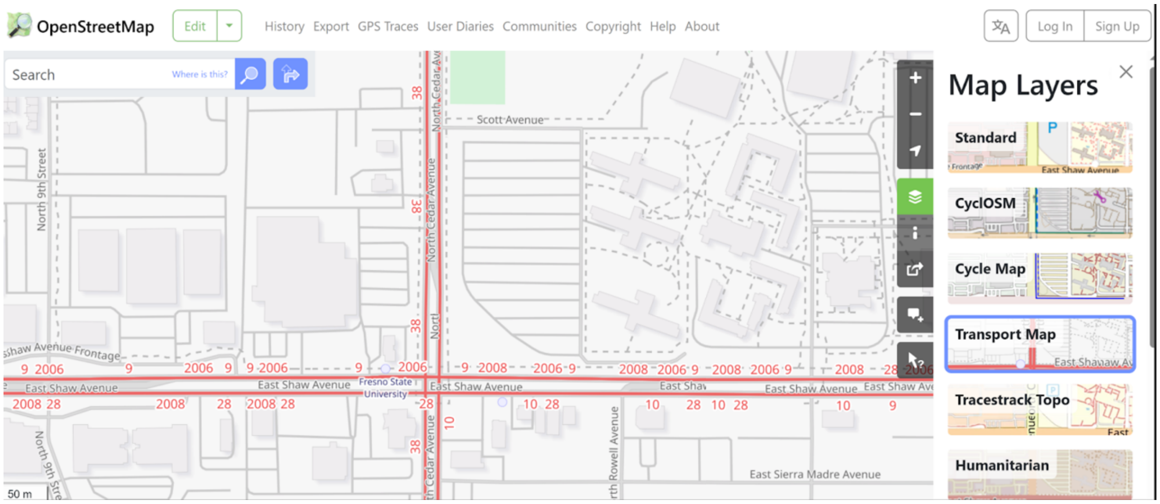



Figure 3 depicts the synthetic Shaw–Cedar traffic dataset used for DT calibration and AI training (2020–2030), capturing high-resolution temporal dynamics and external factors (weather/events).

## TraCI-Based Live Injection

To emulate real-time system operation, synthetic demand is injected into the SUMO simulation using the Traffic Control Interface (TraCI). At each discrete time step, demand inputs are read and translated into vehicle flows dynamically entering the network.

This approach allows the simulation to function as a live Digital Twin engine, where traffic conditions evolve continuously in response to time-varying demand patterns. The integration of synthetic demand with TraCI enables seamless transition to real-time data streams in future deployments without modification to the underlying system architecture.

Figure 4. Synthetic Shaw–Cedar Dataset (2020–2030) for Digital Twin Calibration and AI Training

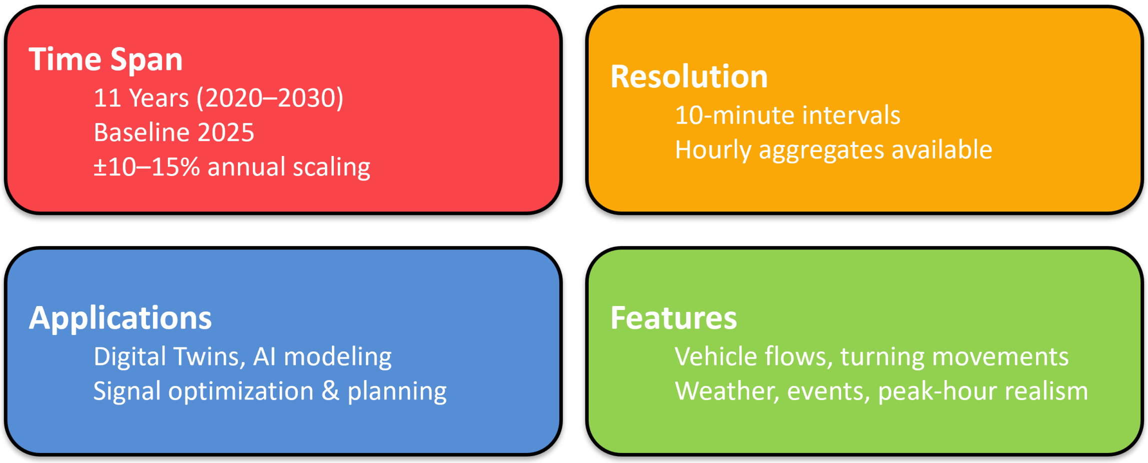



Figure 4 illustrates the OpenStreetMap (OSM) extraction of the Shaw–Cedar intersection used as the base roadway geometry for the Digital Twin.

## Data Acquisition and Preprocessing

SMART-TWIN is designed to ingest traffic information from multiple sources, including the following:

- OpenStreetMap (OSM) roadway geometry
- Synthetic or measured traffic demand
- Lane-level volume and turning movement data
- Contextual information such as time-of-day and weather conditions

All inputs are standardized into a unified schema with a 10-minute temporal resolution. This interval provides sufficient granularity to capture peak-period dynamics and operational disruptions while maintaining computational efficiency.

The framework supports two operational modes:

1. Synthetic mode, used for early deployment and controlled evaluation
2. Live mode, in which real-time sensor data can replace synthetic inputs without changes to the system architecture

## Microscopic Simulation and Live Emulation

The core of the Digital Twin is a microscopic traffic simulation implemented using the Simulation of Urban Mobility (SUMO) platform. Roadway geometry and network topology are derived from OpenStreetMap and refined using standard tools to preserve lane-level connectivity and turning movements.

Traffic demand is injected into the simulation at each time step using the TraCI interface, allowing the model to emulate evolving traffic conditions. This enables the Digital Twin to operate as a live engine that reflects time-varying demand, operational disruptions, and capacity changes such as lane closures or signal failures.

Simulation outputs include lane-level flows, vehicle counts, speeds, delays, and network-level performance measures. These outputs represent the Digital Twin state and are used as inputs to downstream analytics.

## AI-Based Operational State Inference

To move beyond visualization and raw performance monitoring, SMART-TWIN incorporates supervised machine learning to infer operational conditions automatically.

For each time interval, a feature vector is constructed using the following:

- Lane-level traffic volumes and turning movements
- Aggregate intersection demand
- Temporal features such as hour of day and day of week

Machine learning models classify the system into one of four operational states:

- Normal operation
- Stop-and-go conditions (signal failure or control degradation)
- One-lane closure
- Two-lane closure

A Random Forest classifier is used to capture nonlinear relationships among traffic variables and to provide robust disruption detection under varying demand levels. The classification results serve as a trigger for downstream performance and impact analysis.

## Performance Evaluation and Decision Support

SMART-TWIN translates detected operational conditions into interpretable performance measures that support operational and planning decisions.

Two key evaluation components are used:

## Queueing-Based Backlog Estimation

A discrete-time backlog model estimates the accumulation of unmet demand over time, representing the growth of congestion under capacity reductions or control failures.

## Delay-to-Cost Conversion

Average vehicle delay obtained from the simulation is converted into a user-cost metric using value-of-time assumptions. This enables comparison of operational scenarios in terms of economic impact.

Together, these metrics provide a progression from:

Traffic state detection:
- congestion growth
- user delay
- decision-relevant cost impact

This structure allows transportation agencies to evaluate the severity of disruptions, prioritize mitigation actions, and assess the operational benefits of infrastructure or control strategies.

## Corridor-Scale Implementation

The SMART-TWIN framework was demonstrated using a corridor-scale Digital Twin of the Shaw–Cedar intersection in Fresno, California. The network was constructed from OpenStreetMap data and configured to represent lane-level operations at a high-volume signalized intersection serving the Fresno State area.

The modular architecture ensures that the framework can be extended to multiple intersections or corridor networks, with scalability primarily constrained by simulation and data processing resources rather than model structure.

# 5. AI-Driven Data Preparation and Analysis Pipeline

To operationalize the SMART-TWIN framework beyond simulation and visualization, a structured, multi-phase AI-driven analysis pipeline was developed. This pipeline integrates long-horizon traffic data, scenario reconstruction, supervised learning, and performance-based impact assessment to enable actionable insights for transportation operations.

## Phase I: Data Loading and Scenario Preparation

The full 2020–2030 traffic dataset was loaded and validated to ensure temporal consistency, completeness of lane-level turning movements, and correctness of aggregate volumes. Each observation corresponds to a 10-minute interval, capturing time-varying demand and operational conditions at the Shaw–Cedar intersection.

From the baseline dataset, four operational scenarios were reconstructed:

- Normal: Nominal signalized operation with all lanes active
- Stop-and-Go: Emulation of signal failure using stop-controlled behavior
- Lane1 Closed: Partial capacity reduction due to closure of one major through lane
- Lane2 Closed: Severe capacity reduction due to closure of two lanes, including through and turning movements

Scenario generation was performed by systematically modifying lane-level flows while preserving realistic temporal demand patterns. Data integrity checks ensured consistency across lane volumes, total demand, and timestamps. The final dataset was structured for supervised learning by assigning each observation a corresponding scenario label.

## Phase II: AI Model Training and Evaluation

Two supervised machine learning models were developed to classify operational states from traffic observations:

- A multinomial logistic regression model was used as an interpretable baseline
- A random forest classifier was used to capture nonlinear interactions among traffic features

Input features included time-based variables (hour of day, day of week) and lane-level turning movement volumes for all approaches. The dataset was partitioned using a stratified train-test split to maintain class balance.

Model performance was evaluated using classification accuracy and confusion matrices. Feature importance analysis was also conducted for the random forest model to identify the most influential traffic movements contributing to scenario differentiation.

## Phase III: Scenario Impact Analysis

Beyond classification, SMART-TWIN quantifies the operational impacts of each scenario using queueing-based performance modeling.

A backlog simulation framework was used to estimate congestion accumulation under varying capacity conditions. From this analysis, key performance indicators were derived, including the following:

- Queue length evolution
- Average delay per vehicle
- Capacity-to-demand ratios
- Congestion buildup timelines

These metrics enable direct comparison across normal, lane closure, and signal failure scenarios, translating AI-based classifications into interpretable operational impacts.

## Phase IV: Traffic Trend Visualization

To contextualize scenario impacts within long-term demand evolution, multiple visualization approaches were developed:

- A three-dimensional monthly Average Daily Traffic (ADT) surface (2020–2030)
- A two-dimensional yearly ADT trend curve
- A dataset overview infographic
- An intersection movement diagram illustrating directional flow patterns

These visualizations provide insight into temporal traffic variability and support interpretation of model outputs within broader system trends.

Figure 5. Confusion Matrices Comparing Logistic Regression and Random Forest Performance for Traffic State Classification

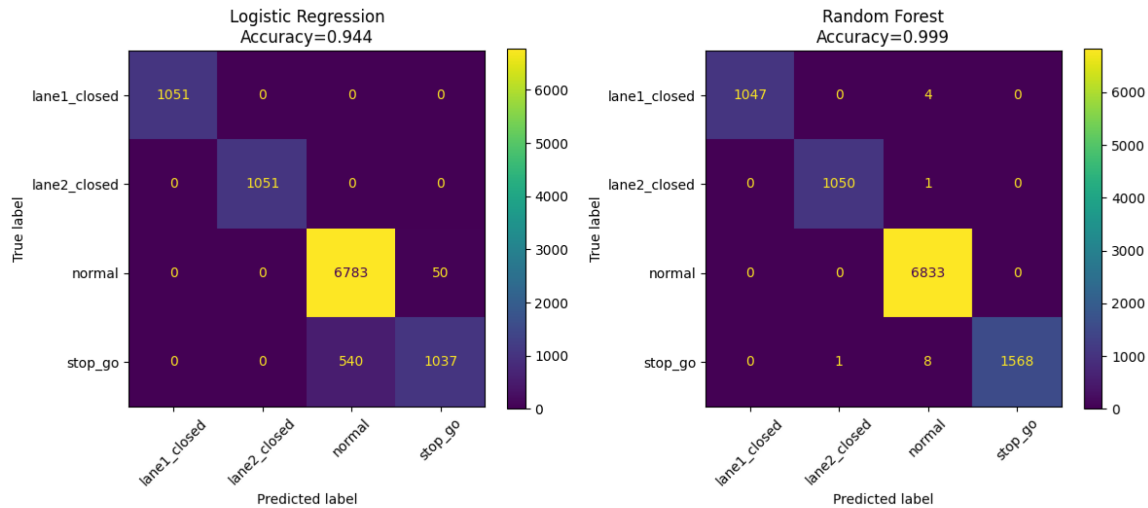



Figure 5 illustrates the confusion matrices for AI-based operational state classification. The logistic regression model achieves 94.4% accuracy, with minor confusion between normal and stop-and-go conditions, while the random forest model achieves near-perfect accuracy (99.9%), demonstrating strong capability in distinguishing complex traffic scenarios.

Figure 6. AI Classification Performance Comparison Between Logistic Regression and Random Forest Models

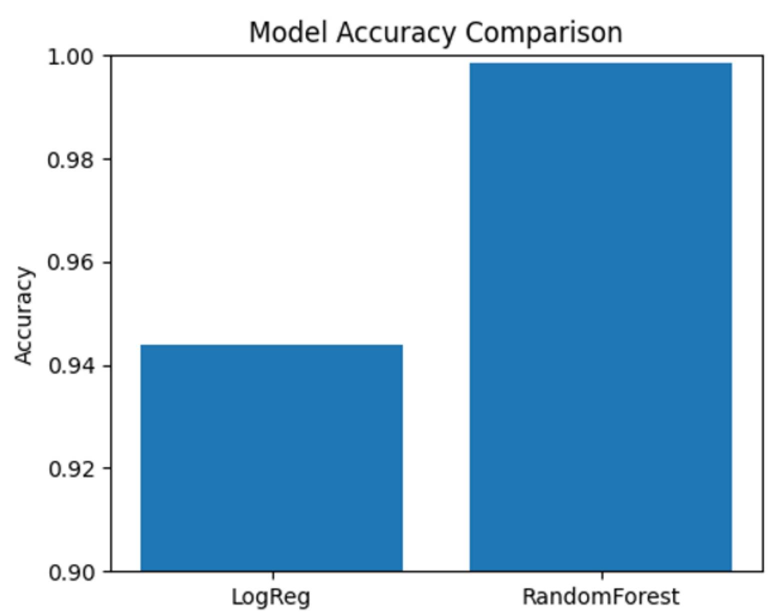



Figure 6 confirms near-perfect operational state detection using the random forest classifier on the Shaw–Cedar dataset.

## Phase V: Visualization and Presentation Integration

All analytical outputs were consolidated into a structured visualization and reporting suite. High-resolution figures were generated and integrated into a consistent scientific layout with standardized labeling, captions, and visual hierarchy.

This integration supports effective dissemination across:

- Technical reports
- Academic publications
- Transportation agency presentations

# 6. Scenario-Based Operational Analysis

To evaluate the practical implications of AI-based state detection, SMART-TWIN was used to analyze traffic behavior under representative disruption scenarios. This section examines how lane closures and signal failures affect intersection performance and congestion dynamics.

## Lane Closure Modeling

Lane closure scenarios were implemented by selectively deactivating lanes and controlling vehicle entry using TraCI. These scenarios simulate partial and severe capacity reductions along major approaches.

Key performance indicators (KPIs) include the following:

- Queue length
- Throughput
- Travel time
- Intersection delay

The analysis enables quantification of congestion buildup and capacity degradation under reduced lane availability.

Figure 7. Baseline SUMO Simulation Under Normal Operating Conditions

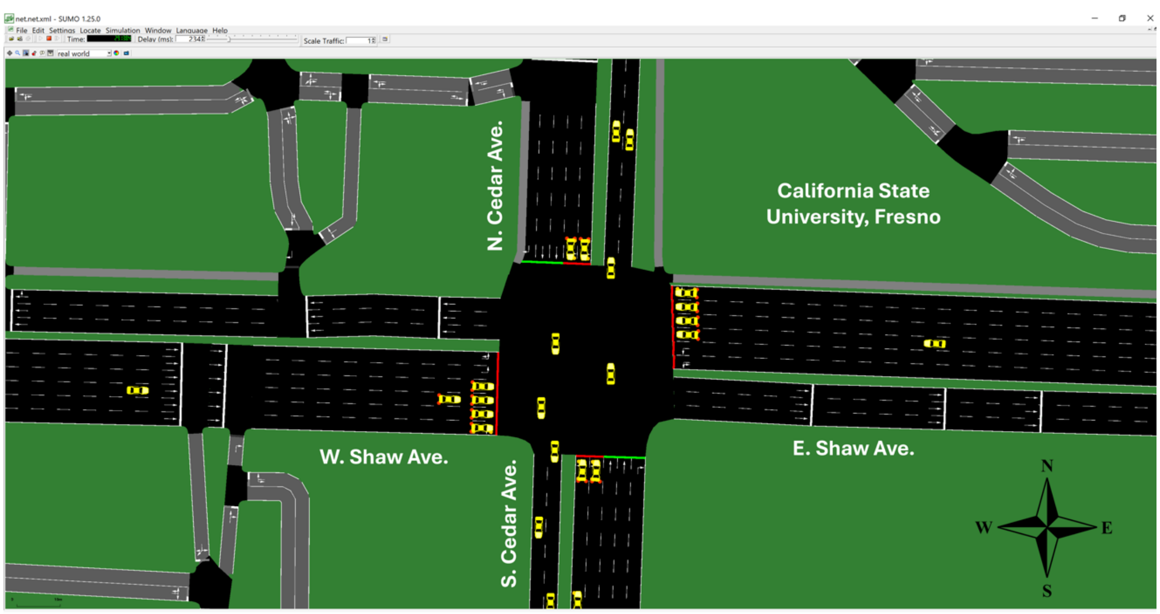


Figure 7 shows the baseline SUMO simulation under normal operation conditions without traffic disruptions.

## Signal Failure Modeling

Signal failure conditions were emulated by replacing signalized control with stop-controlled operation. Both baseline and failure scenarios were simulated to evaluate differences in traffic performance, as shown in Figure 8.

Comparative analysis focuses on the following:

- Queue growth over time
- Delay propagation across approaches
- Intersection throughput degradation

This approach provides insight into the operational impact of signal outages and the resulting traffic instability.

Figure 8. SUMO Simulation of a Two-Lane Closure on Shaw Avenue (Right) Compared with Normal Operation (left)

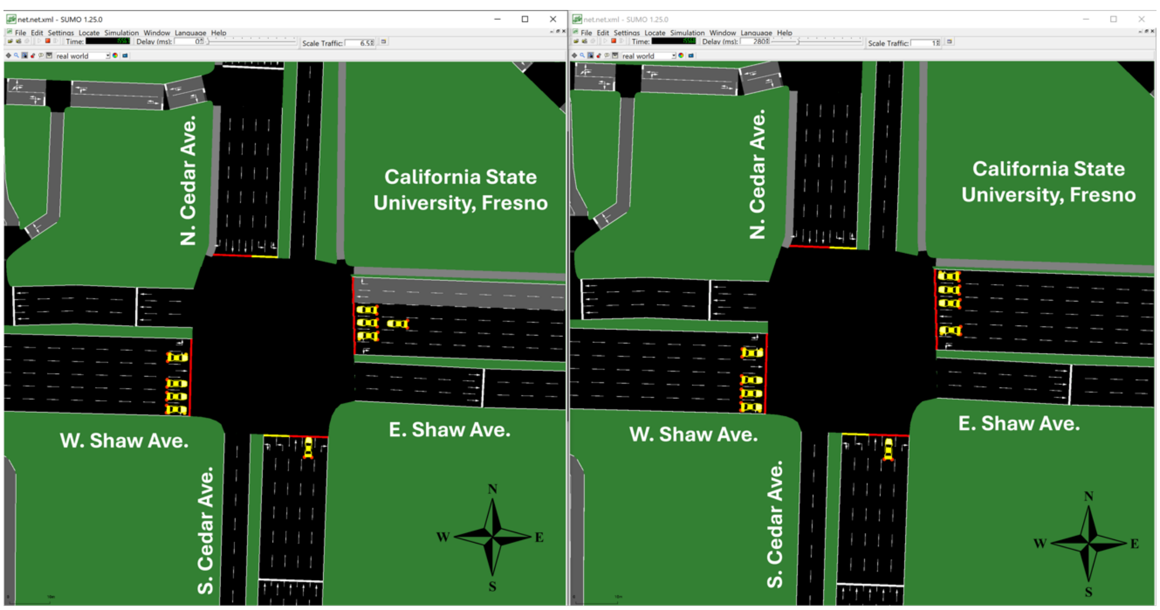



Figure 8 compares normal operation and a two-lane closure scenario at the Shaw–Cedar intersection. The lane closure scenario results in reduced discharge capacity and visible queue buildup along Shaw Avenue, demonstrating how localized capacity loss propagates congestion upstream.

These scenario-based analyses demonstrate how SMART-TWIN translates AI-driven state detection into interpretable operational insights for traffic management and planning.

Figure 9. SUMO Simulation of Stop-and-Go Traffic Due to Signal Failure (Right) Compared with Normal Operation (Left)

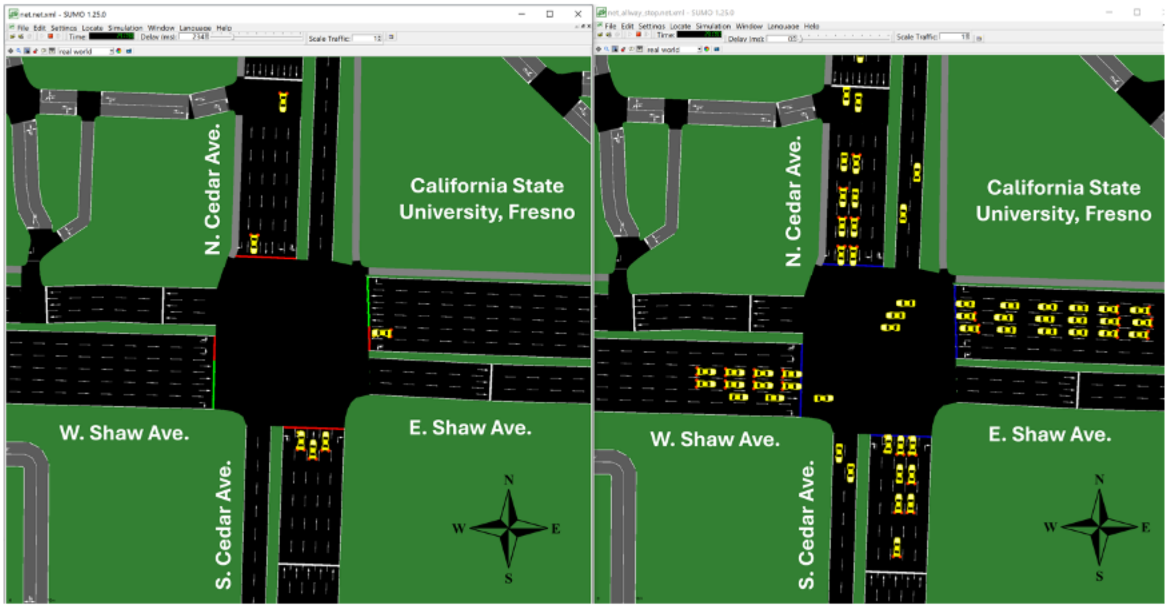



Figure 9 compares normal operation with stop-and-go conditions caused by signal failure.

# 7. Queueing-Based Backlog Proxy and Cost Proxy

To translate AI-based traffic state detection into interpretable operational and economic impacts, SMART-TWIN incorporates queueing-based backlog modeling and delay-to-cost conversion. These proxies enable quantification of congestion severity and user impact under different disruption scenarios.

## Backlog (Queue) Proxy

To quantify congestion buildup, a discrete-time backlog proxy is defined based on arrival and service dynamics.

Let:

- $A_k$: number of arriving vehicles during time interval k
- $S_k$: number of served (discharged) vehicles

The backlog evolves as:

$$B_{k+1} = \max\{0, B_k + A_k - S_k\} \tag{6}$$

where $B_k$ represents the backlog (queue) at time step k, measured in vehicles.

In practice:

- $A_k$ is obtained from injected demand (lane-level volumes)
- $S_k$ is estimated from simulated throughput or saturation-flow discharge

For multi-lane or multi-approach systems, the backlog equation is applied element-wise.

## Normalized Severity Index

To provide a scale-independent measure of congestion, a normalized severity index is defined as:

$$\eta_k = \frac{B_k}{\max(\epsilon, A_k)} \tag{7}$$

where $\epsilon > 0$ avoids division by zero.

This metric reflects the relative imbalance between demand and service capacity and serves as an indicator of the severity of disruption.

Figure 10. SUMO Vehicles-in-Network a 24-hour Period Across Scenarios

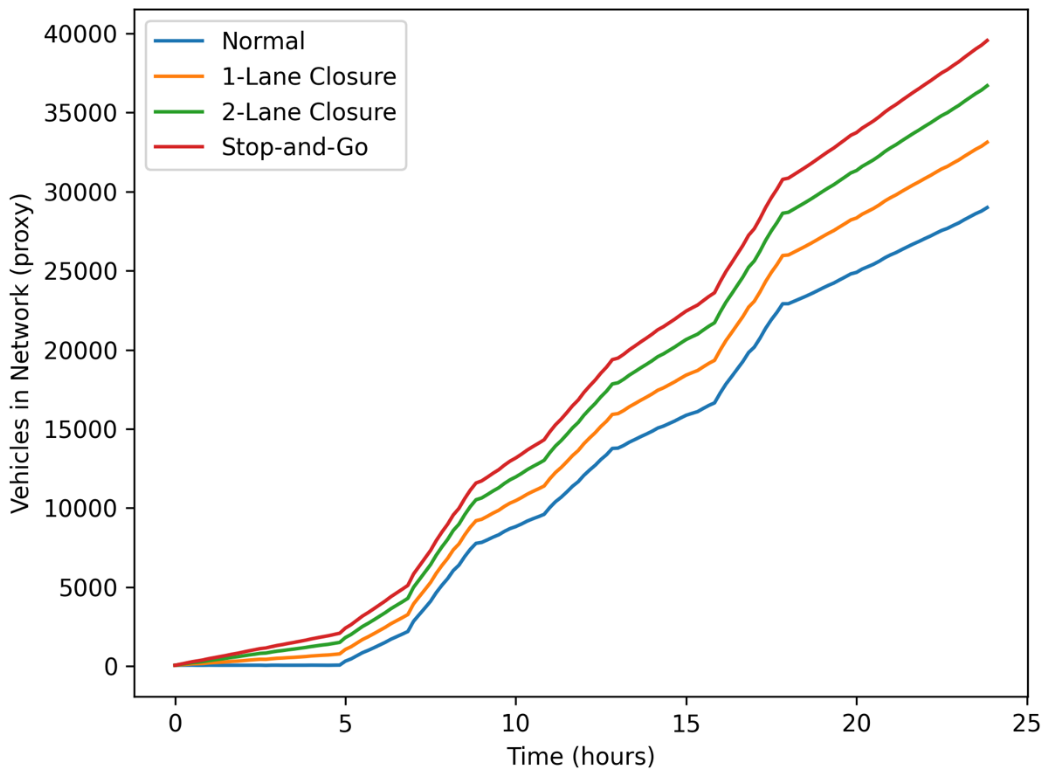



Figure 10 shows the evolution of the backlog proxy across operational scenarios over a representative 24-hour period. The results highlight nonlinear backlog growth, with clear separation between normal operation, lane closures, and stop-and-go conditions as congestion accumulates over time.

## Delay-to-Cost Conversion Using Value of Time (VOT)

To translate delay into an interpretable user-impact metric, a Value of Time (VOT) model is applied.

When average per-vehicle delay is directly available from SUMO (in seconds per vehicle), the daily user-cost proxy for scenario m∈S is computed as:

$$C^{(m)} \approx c_{\mathrm{VOT}} \cdot V \cdot \frac{\bar{d}^{(m)}}{3600} \tag{8}$$

where:

- $V$: average daily demand (vehicles/day)
- $\bar{d}^{(m)}$: mean delay (seconds/vehicle)

- $c_{\text{VOT}}$: value of time ($/vehicle-hour)

### Delay Estimation from Backlog Proxy

When delay is not directly available, it can be approximated from backlog dynamics.

Define the delay proxy (in time steps) as:

$$\delta^{(m)}_{\text{proxy}} = \mathbb{E}\left[\frac{B_k}{\max(\epsilon, A_k)}\right]. \tag{9}$$

Using the simulation time interval Δt (e.g., 600 seconds for 10-minute intervals), the delay is approximated as:

$$\bar{d}^{(m)} = \delta^{(m)}_{\text{proxy}} \Delta t. \tag{10}$$

This estimate is then substituted into Equation (8) to compute the cost proxy.

Figure 11. Average Network Delay Proxy Across Operational Scenarios

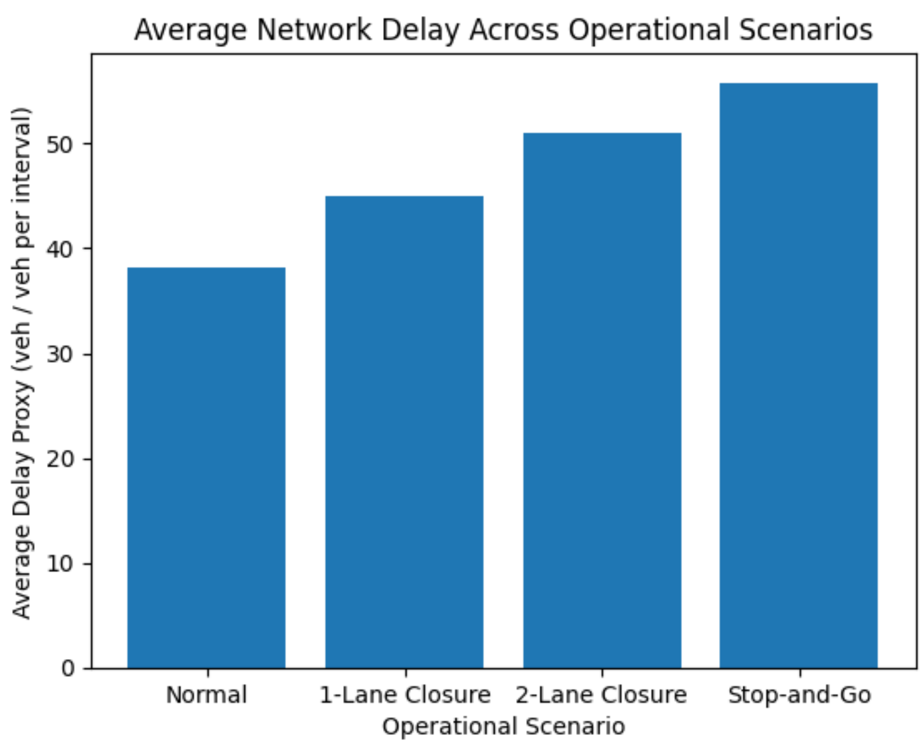
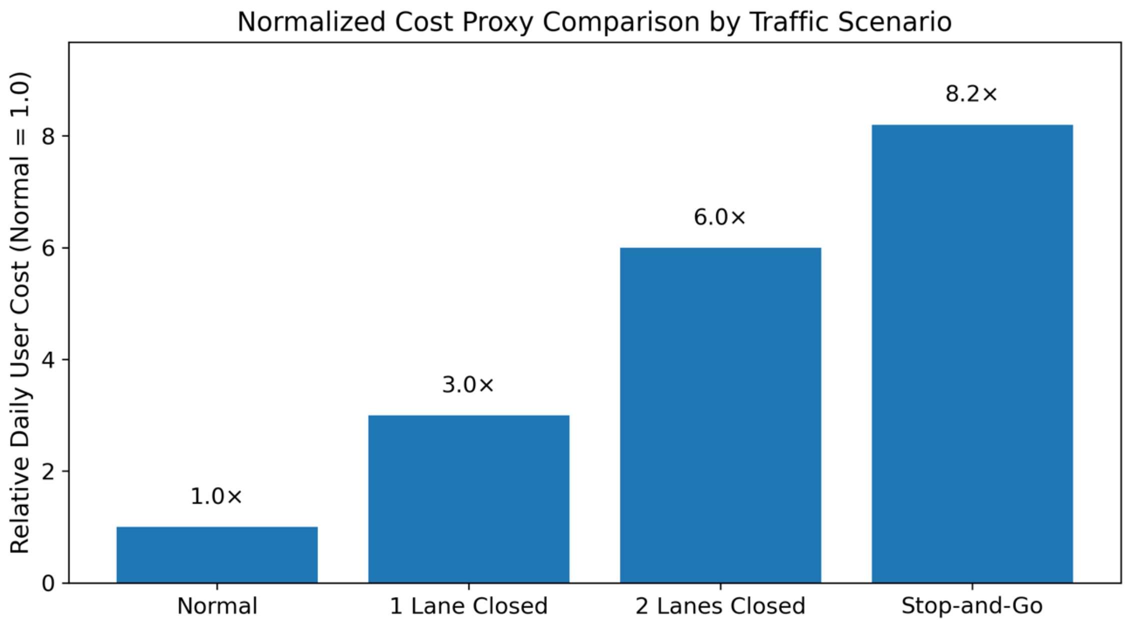



Figure 11 presents the average network delay proxy across operational scenarios. The results indicate nonlinear delay escalation under capacity reductions and signal failures, underscoring the sensitivity of delay to disruptions.

## Cost Proxy Visualization

Figure 12 shows the normalized daily user-cost proxy across scenarios.

Figure 12. Normalized Daily User-Cost Proxy Across Operational Scenarios
(Relative to normal operation)




The figure demonstrates compounding user impact under lane closures and stop-and-go conditions, independent of absolute value-of-time assumptions.

## Integrated Interpretation and Decision Support

Figures 10–12 illustrate a progressive abstraction of system performance:

- Backlog (vehicles) → physical congestion
- Delay (seconds) → operational degradation
- Cost proxy → user-facing impact

This progression enables SMART-TWIN to bridge the gap between simulation outputs and decision-making metrics.

Within a traffic management context, these outputs can be mapped to operational thresholds. For example:

- Sustained lane-closure detection combined with rapid backlog growth can trigger mitigation strategies such as signal retiming, access control, or traveler information updates
- Early detection of stop-and-go conditions due to signal failure can prompt maintenance dispatch or temporary control overrides

In this sense, SMART-TWIN functions as a decision-support layer, enabling proactive traffic management rather than passive monitoring.

# 8. Robustness and Sensitivity Analysis

To evaluate the stability and generalizability of SMART-TWIN under varying operating conditions, a robustness and sensitivity analysis was conducted using controlled perturbations of traffic demand and disruption severity. Rather than introducing new scenarios, this analysis leverages the existing synthetic dataset and reconstructed operational states to assess whether classification performance and operational impact metrics remain consistent under moderate uncertainty.

## Demand Scaling Sensitivity

Traffic demand was uniformly scaled by factors of 0.8×, 1.0×, and 1.2× relative to baseline demand profiles for each operational scenario. These scaling levels represent realistic variations arising from seasonal changes, special events, or forecasting uncertainty. For each demand level, the Digital Twin simulation, AI-based operational state classification, and queueing-based performance metrics were recomputed.

Across all demand levels, the Random Forest classifier maintained consistently high accuracy, with no observable degradation in detecting lane closures or Stop-and-Go conditions. This indicates that the learned decision boundaries are robust to moderate variations in demand magnitude.

In contrast, the normalized cost proxy exhibited nonlinear growth as demand increased, particularly under disruption scenarios. While normal operation scaled approximately linearly with demand, lane closure and signal failure scenarios showed disproportionate increases in delay and backlog accumulation, reflecting compounding congestion effects under reduced capacity.

These results highlight a key property of SMART-TWIN: state recognition remains stable, while impact metrics appropriately reflect system stress, enabling reliable detection under uncertainty without masking operational severity.

## Disruption Severity Sensitivity

In addition to demand scaling, sensitivity to disruption severity was evaluated by varying the degree of capacity reduction (e.g., partial versus full lane closures).

As expected:

- Increased severity led to faster backlog accumulation
- Delay and cost proxies exhibited monotonic growth with capacity loss

- System degradation patterns remained consistent across severity levels

Importantly, the AI classifier correctly identified the operational state across all severity variations, indicating that classification performance is not dependent on a specific disruption magnitude.

Overall, these findings demonstrate that SMART-TWIN is robust to moderate modeling uncertainty and preserves both:

- High classification accuracy
- Reliable, decision-relevant performance trends

# 9. Early-Warning Detection Capability

Beyond static classification, SMART-TWIN is designed to detect emerging disruptions **before severe congestion develops**, enabling proactive intervention.

This capability was evaluated by measuring **detection latency** following transitions from normal operation to disruption scenarios. For each lane closure and Stop-and-Go signal failure case, disruptions were introduced at known simulation times.

Detection latency was defined as the number of **10-minute intervals** required for the AI classifier to consistently identify the correct operational state.

## Key Observations

- Lane closures were detected within 1–2 intervals, reflecting gradual changes in lane utilization and queue formation
- Stop-and-Go signal failures were detected within a single interval in most cases, due to abrupt breakdown in discharge behavior

In all cases, detection occurred prior to peak congestion, indicating that SMART-TWIN provides actionable early signals rather than retrospective diagnostics.

This early-warning capability enables:

- Proactive signal retiming
- Rapid incident response
- Timely traveler information dissemination

# 10. Role of AI-Based State Classification in SMART-TWIN

Within the SMART-TWIN framework, AI-based operational state classification serves as a decision-triggering layer, rather than an end objective.

Using the 2020–2030 Shaw–Cedar dataset, the Random Forest classifier reliably identifies key operational states, including:

- Lane closures
- Stop-and-Go conditions due to signal failure

These classifications act as inputs to downstream modules, including:

- Queueing-based backlog estimation
- Delay-to-cost conversion

By explicitly separating:

- State recognition (AI layer)
- Performance evaluation (queueing + cost models)

SMART-TWIN ensures that operational decisions are both:

- Robustly triggered, and
- Quantitatively justified

Scalability Considerations

The computational structure of SMART-TWIN supports scalable deployment:

- AI inference and queueing-based computations scale linearly with the number of approaches/intersections
- The primary computational bottleneck is microscopic simulation (SUMO + TraCI) rather than classification

This distinction is critical, as it implies that system-wide deployment is feasible without significant AI overhead, with optimization efforts focused on simulation efficiency.

This study relies on synthetic demand and simulated environments; while this enables controlled evaluation, real-world deployment may introduce additional uncertainties such as sensor noise, driver variability, and incomplete data. These factors will be addressed in future field validation.

# 11. Conclusion

This study developed and demonstrated SMART-TWIN, an Artificial Intelligence (AI)–driven Digital Twin framework for traffic operations monitoring, disruption detection, and performance-based decision support. The framework integrates microscopic traffic simulation, long-horizon synthetic demand generation, supervised machine learning, and queueing-based performance modeling within a unified and scalable architecture.

The corridor-scale case study at the Shaw–Cedar intersection in Fresno, California, demonstrated the ability of SMART-TWIN to emulate time-varying traffic conditions, accurately identify operational disruptions, and quantify their impacts in terms of congestion growth, delay, and user cost. The machine learning model achieved consistently high classification accuracy, while disruption events were detected within one to two analysis intervals, enabling timely response. Scenario-based results further revealed that capacity reductions and control failures lead to nonlinear increases in congestion and economic impact, underscoring the importance of early detection and proactive traffic management.

A key contribution of this work is the integration of robust operational state recognition with interpretable performance and cost metrics, allowing transportation agencies to move beyond passive monitoring toward actionable, data-driven decision-making. This capability directly supports applications such as incident response, work zone planning, operational optimization, and infrastructure investment prioritization.

The study also demonstrated the effectiveness of using long-horizon synthetic demand to support early Digital Twin deployment in environments where continuous high-resolution field data are not yet available. This approach enables controlled system evaluation and model development while maintaining compatibility with future real-time data integration.

The SMART-TWIN architecture is inherently modular and scalable, supporting extension from single intersections to corridor-level and network-wide deployments. Future work will focus on integrating real-time sensor and signal data, conducting field validation, improving robustness under real-world uncertainty, and incorporating adaptive control strategies to enable closed-loop operational optimization.

Overall, the results indicate that AI-enabled Digital Twins provide a practical and scalable foundation for next-generation transportation systems, enabling proactive operations, improved system reliability, and data-driven infrastructure decision-making.

# Abbreviations and Acronyms

| | |
|---|---|
| AI | Artificial Intelligence |
| DT | Digital Twin |
| SMART-TWIN | Scalable Monitoring and AI-Driven Real-Time Digital Twin |
| ML | Machine Learning |
| RF | Random Forest |
| SUMO | Simulation of Urban Mobility |
| TraCI | Traffic Control Interface |
| OSM | OpenStreetMap |
| JOSM | Java OpenStreetMap Editor |
| KPI | Key Performance Indicator |
| VOT | Value of Time |
| ITS | Intelligent Transportation Systems |
| DoT | Department of Transportation |

# Bibliography


Ali, W. A., Roccotelli, M., & Fanti, M. P. 2023. *Digital twin in intelligent transportation systems: A Review.* echarge4drivers. https://echarge4drivers.eu/wp-content/uploads/2023/07/Digital-Twin-in-Intelligent-Transportation-Systems-a-Review.pdf

Federal Highway Administration (FHWA). 2021. *Traffic Incident Management (TIM) program.* U.S. Department of Transportation.

Ge, C., et al. 2024. *Digital Twin Intelligent Transportation System (DT-ITS): A systematic literature review and roadmap.* IET Intelligent Transport Systems. https://digital-library.theiet.org/doi/full/10.1049/itr2.12539

Gopalakrishnan, K., and Peeta, S. 2021. Applications of digital twins in intelligent transportation systems. *Journal of Transportation Engineering*, *147*(10).

Kusić, K., et al. 2023. A digital twin in transportation: Real-time synergy of traffic simulation and real-world traffic data. *Advanced Engineering Informatics*. https://www.sciencedirect.com/science/article/pii/S1474034622003160

Kusić, K., and Schumann, A. 2022. Building a motorway Digital Twin in SUMO: Real-time microscopic simulation coupled with traffic data. https://arodes.hes-so.ch/record/14883/files/Schumann%202022%20Building%20motorway.pdf

Llagostera-Brugarola, E., et al. 2025. A digital twin for intelligent transportation systems in highways. *Applied Sciences*. https://www.mdpi.com/2076-3417/15/13/7454

Lopez, P. A., et al. 2018. Microscopic traffic simulation using SUMO. In *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC)*.

Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F.-Y. 2015. Traffic flow prediction with big data: A deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, *16*(2), 865–873.

Pradhananga, R., et al. 2025. Digital Twin-aided municipal traffic control. *SUMO User Conference (preprint).* https://www.tib-op.org/ojs/index.php/scp/article/view/2619

Rundel, S., et al. 2023. Leveraging digital twin and game-engine for traffic simulations and visualizations. *Frontiers in Virtual Reality*. https://www.frontiersin.org/articles/10.3389/frvir.2023.1048753

Smart Nation Singapore. 2018. *Virtual Singapore Smart City Project.*

Xing, L., et al. 2025. Advancing intelligent transportation through digital twin: Challenges, models, and future prospects. *Ad Hoc Networks*. https://www.sciencedirect.com/science/article/pii/S1570870525003257

# About the Author

**Hovannes Kulhandjian**

Dr. Hovannes Kulhandjian is a tenured full Professor in the Department of Electrical and Computer Engineering at California State University, Fresno (Fresno State). He joined Fresno State in Fall 2015 as a tenure-track faculty member. Before that, he was an Associate Research Engineer in the Department of Electrical and Computer Engineering at Northeastern University. He received his B.S. in Electronics Engineering with high honors from the American University in Cairo (AUC) in 2008 and his M.S. and Ph.D. degrees in Electrical Engineering from the State University of New York at Buffalo in 2010 and 2014, respectively. His current research interests are in applied machine learning, autonomous vehicle navigation, wireless communications, and networking, with applications to underwater and visible light communications and networking geared towards Intelligent Transportation Systems (ITS).

Dr. Kulhandjian has received eight research grants from the Fresno State Transportation Institute (FSTI), the Claude C. Laval III Award for Commercialization of Research, Innovation, and Creativity in 2025 and 2021, and the Claude C. Laval Award for Innovative Technology and Research 2020 at Fresno State. In April 2021, as a PI, he received a grant from the Department of Defense (DOD) Research and Education Program for Historically Black Colleges and Universities and Minority-Serving Institutions (HBCU/MSIs) Equipment/Instrumentation to establish a "Secure Communications and Embedded Systems Laboratory at California State University, Fresno".

Dr. Kulhandjian is an active member of the Association for Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE). He is a Senior Member of IEEE. He is serving as a Guest Editor for the Special Issue "Advances in Intelligent Transportation Systems (ITS)". He has served as a Guest Editor for the IEEE Access Special Section Journal, Session Co-Chair for the IEEE UComms'20 Conference, Session Chair for the ACM WUWNet'19 Conference, and Publicity Co-Chair for the IEEE BlackSeaCom Conference. He also serves as a member of the Technical Program Committee (TPC) for ACM and IEEE Conferences such as GLOBECOM 2022, WTS 2022, WD 2021, WD 2018, ICC 2018, WUWNet 2020, and WiMob2019. He is a recipient of the Outstanding Reviewer Award from ELSEVIER Ad Hoc Networks and ELSEVIER Computer Networks.

## MTI FOUNDER

### Hon. Norman Y. Mineta



MINETA TRANSPORTATION INSTITUTE