RESEARCH PROPOSAL

A Simulation Framework for Evaluating AI-Powered Traffic Control

Submitted by:

Kidenge Elisha Odhiambo

Reg. No.: SM24/11854/24

Program: Master of Science in Computer Science

Supervisors:

Dr. Zachary Bosire, PhD

Chair, Computer Science Department,

Egerton University

Dr. Duke Oeba, PhD

Lecturer and Researcher - Renewable Energy

Department of Physics, Egerton University

Department of Computer Science

Egerton University

Nakuru, Kenya

September 27, 2025

Declaration

I, Kidenge Elisha Odhiambo, Registration Number SM24/11854/24, hereby declare that this research proposal titled ”A Simulation Framework for Evaluating AI-Powered Traffic Control in Kenyan Urban Environments” is my original work and has not been previously submitted for any academic award in this or any other institution. All sources of information used have been duly acknowledged through complete referencing.

Kidenge Elisha Odhiambo

Registration Number: SM24/11854/24

Department of Computer Science

Egerton University

Dedication

This research proposal is dedicated to the urban commuters of Kenya who endure daily traffic congestion, and to the researchers and policymakers working towards sustainable urban mobility solutions in developing nations.

Acknowledgements

I express my sincere gratitude to my supervisors, Dr. Zachary Bosire and Dr. Duke Oeba, for their invaluable guidance, mentorship, and constructive feedback throughout the development of this research proposal. Their expertise in computational methods and urban systems has been instrumental in shaping this work.

I extend my appreciation to the Department of Computer Science at Egerton University for providing the academic environment and resources necessary for this research. Special thanks to my colleagues and the technical staff for their support and insightful discussions.

I am grateful to my family for their unwavering support and encouragement throughout my academic journey. Their belief in my capabilities has been a constant source of motivation.

Finally, I acknowledge the open-source community whose tools and libraries, particularly SUMO, TensorFlow, and OpenCV, make research of this nature feasible and reproducible.

Abstract

Urban traffic congestion presents significant socio-economic and environmental challenges in rapidly growing cities across Kenya, particularly Nairobi and Nakuru. The existing traffic control infrastructure in these urban centers remains largely inadequate, characterized by static signal timing, limited adaptive capabilities, and insufficient data-driven decision-making processes. This research proposal addresses these challenges by designing and validating a comprehensive simulation framework for evaluating artificial intelligence (AI)-powered traffic control strategies tailored specifically to Kenyan urban environments.

The proposed methodology integrates four core modules within a digital twin environment: (1) synthetic data generation to overcome the scarcity of annotated traffic datasets, (2) a perception module for vehicle detection and tracking using deep learning approaches, (3) a prediction module for short-term traffic flow forecasting employing time-series analysis techniques, and (4) an adaptive control module utilizing reinforcement learning for dynamic traffic signal optimization. The framework will be implemented using open-source simulation tools, primarily the Simulation of Urban Mobility (SUMO), ensuring accessibility and reproducibility.

This research aims to contribute to intelligent transportation systems (ITS) literature by providing an empirically validated simulation framework that addresses the unique challenges of mixed-traffic environments in developing cities. The expected outcomes include an open-source evaluation framework, performance benchmarks comparing AI-driven approaches with conventional methods, and a pathway for future real-world deployments. By leveraging simulation-based methodologies, this research circumvents the substantial infrastructural investments typically associated with ITS implementations while providing a robust testing environment for innovative traffic management strategies.

The significance of this study lies in its potential to inform traffic management policies in Kenya and similar contexts, demonstrating how simulation-driven approaches can bridge the gap between theoretical research and practical implementation in resource-constrained environments.

Keywords: Traffic Simulation, Artificial Intelligence, Intelligent Transportation Systems, Digital Twin, Reinforcement Learning, Urban Mobility, Kenya

Contents

Declaration **i**

Dedication **ii**

Acknowledgements **iii**

Abstract **iv**

List of Algorithms **xi**

List of Abbreviations **xii**

1 Introduction **1**

1.1 Background and Context 1

1.2 Problem Statement 2

1.3 Research Objectives 3

1.3.1 General Objective 3

1.3.2 Specific Objectives 3

1.4 Research Questions 4

1.5 Research Hypotheses 4

1.6 Significance of the Study 5

1.7 Scope and Limitations 6

1.7.1 Scope of the Study 6

1.7.2 Limitations 6

1.7.3 Delimitations 7

1.8 Thesis Structure 7

2 Literature Review **8**

2.1 Introduction to Urban Traffic Management 8

2.2 Traffic Congestion in Developing Cities 8

2.3 Intelligent Transportation Systems (ITS) 9

2.4 Simulation in Transportation Research 10

2.5 Artificial Intelligence in Traffic Management . . . 11
2.5.1 Computer Vision for Traffic Perception . . . 11
2.5.2 Predictive Modeling for Traffic Flow . . . 12
2.5.3 Reinforcement Learning for Traffic Control . . . 12
2.5.4 Multi-Agent Systems for Distributed Control . . . 12
2.6 Digital Twins in Transportation . . . 13
2.7 Traffic Management in the Kenyan Context . . . 14
2.8 Research Gaps and Contribution . . . 14

3 Theoretical Framework **16**
3.1 Introduction . . . 16
3.2 Complex Systems Theory in Transportation . . . 16
3.3 Queueing Theory and Traffic Flow Modeling . . . 17
3.4 Machine Learning Foundations . . . 17
3.4.1 Statistical Learning Theory . . . 17
3.4.2 Deep Learning Theory . . . 18
3.4.3 Reinforcement Learning Theory . . . 18
3.5 Simulation Theory and Methodology . . . 19
3.6 Integration Framework . . . 19

4 Research Methodology **21**
4.1 Research Design . . . 21
4.2 Simulation Environment Setup . . . 21
4.2.1 Network Modeling . . . 22
4.2.2 Simulation Scenarios . . . 22
4.3 Data Generation and Management . . . 22
4.3.1 Synthetic Data Generation . . . 23
4.3.2 Data Quality Assurance . . . 23
4.4 AI Module Development . . . 24
4.4.1 Perception Module . . . 24
4.4.2 Prediction Module . . . 24
4.4.3 Control Module . . . 25
4.5 Integration Framework . . . 26

4.6 Evaluation Methodology 27
4.6.1 Performance Metrics 27
4.6.2 Statistical Analysis 28
4.7 Ethical Considerations 28
4.8 Limitations of the Methodology 28

5 Implementation Plan **30**
5.1 Technical Architecture 30
5.2 Development Environment 30
5.3 Implementation Timeline 31
5.4 Code Structure and Organization 33
5.5 Key Implementation Details 33
5.5.1 SUMO Integration 33
5.5.2 Perception Module Implementation 34
5.5.3 Reinforcement Learning Controller 35
5.6 Testing Strategy 37
5.7 Documentation and Reproducibility 38

6 Expected Results and Evaluation **39**
6.1 Expected Outcomes 39
6.1.1 Technical Outcomes 39
6.1.2 Academic Outcomes 39
6.1.3 Practical Outcomes 40
6.2 Evaluation Framework 40
6.2.1 Performance Metrics 40
6.2.2 Comparative Analysis 41
6.2.3 Statistical Evaluation 42
6.3 Expected Performance Results 42
6.4 Sensitivity Analysis 43
6.5 Validation Approach 43
6.6 Limitations and Mitigations 44

7 Work Plan and Conclusion **45**

7.1 Detailed Work Plan 45

7.2 Resource Requirements 48

7.2.1 Computational Resources 48

7.2.2 Data Resources 48

7.2.3 Human Resources 48

7.3 Risk Assessment and Mitigation 49

7.4 Dissemination Plan 50

7.4.1 Academic Dissemination 50

7.4.2 Practical Dissemination 50

7.4.3 Public Dissemination 51

7.5 Conclusion 51

A SUMO Configuration Example **55**

B Vehicle Type Definitions **56**

C Evaluation Metrics Calculation **57**

C.1 Efficiency Metrics 57

C.1.1 Average Delay per Vehicle 57

C.1.2 Throughput 57

C.2 Environmental Metrics 58

C.2.1 Fuel Consumption 58

C.3 Safety Metrics 58

C.3.1 Time-to-Collision 58

List of Figures

3.1 Theoretical Integration Framework 19

4.1 Research Design Process 21
4.2 Perception Module Architecture 24
4.3 Framework Integration Architecture 27

5.1 System Architecture Overview 30

List of Tables

5.1 Implementation Timeline 33

6.1 Performance Metrics for Evaluation 41

7.1 Detailed Work Plan 48

List of Algorithms

1. Synthetic Traffic Data Generation Algorithm
2. Vehicle Detection and Tracking Algorithm
3. Traffic Flow Prediction using LSTM
4. Reinforcement Learning for Adaptive Traffic Control
5. Performance Evaluation Metrics Calculation

List of Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
CNN	Convolutional Neural Network
CPU	Central Processing Unit
DQN	Deep Q-Network
GIS	Geographic Information System
GPU	Graphics Processing Unit
ITS	Intelligent Transportation Systems
LSTM	Long Short-Term Memory
ML	Machine Learning
MSE	Mean Squared Error
RL	Reinforcement Learning
RMSE	Root Mean Square Error
SUMO	Simulation of Urban Mobility
SVM	Support Vector Machine
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle
YOLO	You Only Look Once

Chapter 1

Introduction

1.1 Background and Context

Urbanization in Kenya has progressed at an unprecedented rate over the past two decades, with current estimates indicating that approximately 30% of the country's population resides in urban areas, a figure projected to reach 50% by 2050 (of Statistics, 2019). This rapid urban growth, particularly evident in cities like Nairobi and Nakuru, has been accompanied by a dramatic increase in vehicle ownership and usage. The number of registered vehicles in Kenya grew from approximately 1.2 million in 2008 to over 3.5 million in 2022, representing an annual growth rate of nearly 8% (Transport and Authority, 2022).

The convergence of population concentration and increased vehicular activity has resulted in severe traffic congestion that now characterizes daily life in major Kenyan urban centers. Nairobi, the capital city, consistently ranks among cities with the worst traffic congestion globally, with commuters spending an average of 60-90 minutes in traffic during peak hours (Otieno et al., 2021). The economic implications are substantial, with estimates suggesting that traffic congestion costs the Kenyan economy approximately KES 100 billion annually in lost productivity, fuel consumption, and environmental damage (Kamau and Ochieng, 2020).

Traditional traffic management approaches in Kenya have primarily relied on static traffic signal systems, manual police direction, and limited surveillance infrastructure. These methods have proven inadequate for addressing the dynamic and complex nature of urban traffic flow, particularly in environments characterized by heterogeneous vehicle types (personal cars, public service vehicles, motorcycles, pedestrians) and often unpredictable driver behavior (Maitha and Nyong'o, 2020). The limitations of existing systems are further exacerbated by infrastructural constraints, limited funding for transportation infrastructure, and bureaucratic hurdles that impede rapid implementation of innovative solutions.

Globally, Intelligent Transportation Systems (ITS) have emerged as a promising approach to addressing urban mobility challenges. These systems leverage advanced technologies including sensors, communication networks, and computational intelligence to optimize traffic flow, enhance safety, and reduce environmental impacts (Zhang et al., 2018). The integration of artificial intelligence (AI) techniques, particularly machine learning and deep reinforcement learning, has demonstrated significant potential in developing adaptive traffic control strategies that can respond dynamically to changing traffic conditions (Chu et al., 2019; Li et al., 2016).

However, the implementation of AI-powered traffic control systems in developing contexts like Kenya faces unique challenges. These include data scarcity for training and validation, infrastructural limitations, computational resource constraints, and the need for solutions that are specifically tailored to local traffic patterns and behaviors (Mohan and Tiwari, 2017). Simulation-based approaches offer a viable pathway for overcoming these challenges by providing controlled environments for developing, testing, and validating traffic management strategies before real-world deployment (Cheng et al., 2020).

1.2 Problem Statement

The core problem addressed by this research is the inadequacy of existing traffic control systems in Kenyan urban environments to effectively manage congestion and optimize traffic flow. This problem manifests through several interconnected challenges:

Data Scarcity and Quality: The development of effective AI-powered traffic management systems requires extensive, high-quality datasets for training, validation, and testing. In Kenya, such datasets are largely unavailable due to limited sensor infrastructure, inconsistent data collection practices, and privacy concerns (Otieno et al., 2021). This data gap impedes the development and validation of data-driven traffic management solutions tailored to local conditions.

Infrastructural Limitations: The implementation of sophisticated ITS infrastructure faces significant barriers in the Kenyan context, including high costs, maintenance challenges, and compatibility issues with existing systems (Maitha and Nyong'o, 2020). The heterogeneous nature of Kenyan traffic, comprising various vehicle types with diverse operational characteristics, further complicates the implementation of standardized solutions.

Behavioral Complexity: Traffic flow in Kenyan urban environments is characterized by unique behavioral patterns, including informal public transport operations, high pedestrian volumes, and often non-compliant driving behaviors (Kamau and Ochieng, 2020). These factors create traffic dynamics that may not be adequately addressed by control strategies developed for more standardized traffic environments.

Evaluation Challenges: The high stakes associated with real-world traffic system modifications necessitate rigorous testing and evaluation before implementation. However, the absence of robust evaluation frameworks specifically designed for Kenyan conditions makes it difficult to assess the potential effectiveness and impacts of proposed solutions (Mohan and Tiwari, 2017).

These challenges collectively contribute to a significant research gap: the lack of a comprehensive, simulation-based framework for developing and evaluating AI-powered traffic control strategies that are specifically tailored to the unique characteristics of Kenyan urban environments. This research seeks to address this gap by proposing a simulation framework that enables the development, testing, and validation of adaptive traffic control strategies in a risk-free virtual environment before potential real-world deployment.

1.3 Research Objectives

1.3.1 General Objective

The general objective of this research is to design, implement, and validate a comprehensive simulation framework for evaluating AI-powered traffic control strategies in Kenyan urban environments, with specific focus on addressing the unique challenges of mixed-traffic conditions, data scarcity, and infrastructural constraints.

1.3.2 Specific Objectives

To achieve the general objective, the following specific objectives have been formulated:

1. To develop a digital twin of a representative Kenyan urban intersection using open-source simulation tools, accurately modeling vehicle types, driver behaviors, and traffic patterns specific to the Kenyan context.

2. To design and implement a synthetic data generation pipeline that produces realistic traffic datasets, addressing the challenge of data scarcity for training and validation of AI models.

3. To develop a perception module based on deep learning techniques for accurate vehicle detection, classification, and tracking within the simulated environment.

4. To design and implement a prediction module utilizing time-series analysis and machine learning approaches for short-term traffic flow forecasting.

5. To create an adaptive traffic control module using reinforcement learning algorithms that dynamically optimizes traffic signal timing based on real-time traffic conditions.

6. To establish a comprehensive evaluation framework with metrics for assessing the performance of AI-powered traffic control strategies against conventional methods.

7. To analyze the feasibility, scalability, and potential implementation pathways for simulation-validated traffic control strategies in real-world Kenyan contexts.

1.4 Research Questions

This research is guided by the following research questions:

1. How can simulation-based approaches effectively overcome data scarcity challenges in developing AI-powered traffic management solutions for Kenyan urban environments?

2. What is the comparative performance of AI-driven adaptive traffic control strategies versus conventional fixed-time and actuated control systems in simulated Kenyan traffic conditions?

3. Which AI techniques and architectures are most effective for the specific challenges of mixed-traffic environments characteristic of Kenyan cities?

4. How can synthetic data generation techniques be optimized to produce training datasets that accurately represent the complexity and variability of real-world Kenyan traffic?

5. What are the key implementation considerations and potential barriers for deploying simulation-validated AI traffic control strategies in real-world Kenyan contexts?

6. How do environmental factors, such as weather conditions and special events, impact the performance of AI-powered traffic control systems in simulated Kenyan urban environments?

1.5 Research Hypotheses

Based on the research objectives and questions, the following hypotheses are proposed:

1. **H1:** AI-powered adaptive traffic control strategies will demonstrate statistically significant improvements in key performance metrics (average waiting time, throughput, emissions) compared to conventional traffic control methods in simulated Kenyan urban environments.

2. **H2:** Synthetic data generation techniques can produce training datasets that enable the development of AI models with performance comparable to models trained on real-world data for traffic management applications.

3. **H3:** Reinforcement learning-based traffic control algorithms will exhibit better adaptation to the heterogeneous and often unpredictable traffic patterns characteristic of Kenyan cities compared to rule-based adaptive systems.

4. **H4:** The proposed simulation framework will provide a cost-effective and scalable approach for evaluating traffic management strategies, reducing the need for expensive pilot deployments in the initial stages of development.

1.6 Significance of the Study

This research offers several significant contributions to both academic knowledge and practical applications in urban traffic management:

Theoretical Contributions:

- Advances the theoretical understanding of how AI techniques can be adapted for traffic management in developing urban contexts with unique characteristics and constraints.
- Contributes to the field of digital twin applications in transportation by demonstrating their utility in resource-constrained environments.
- Provides insights into the transferability of traffic management strategies developed in simulation environments to real-world contexts with different operational conditions.

Practical Implications:

- Offers a cost-effective framework for developing and testing traffic management strategies without the risks and expenses associated with real-world implementations.
- Provides Kenyan traffic authorities with evidence-based insights into the potential benefits of AI-powered traffic control systems.

- Establishes a foundation for future research and development in intelligent transportation systems tailored to African urban contexts.

Methodological Innovations:

- Develops novel approaches for synthetic data generation that address the specific challenges of data scarcity in developing regions.
- Creates integrated evaluation metrics that capture the multi-dimensional impacts of traffic management strategies (efficiency, environmental, safety).
- Demonstrates how open-source tools can be leveraged to create sophisticated simulation environments accessible to researchers in resource-constrained settings.

1.7 Scope and Limitations

1.7.1 Scope of the Study

This research focuses on the development and validation of a simulation framework for evaluating AI-powered traffic control strategies. The scope encompasses:

- Simulation of a single representative intersection in a Kenyan urban environment, with potential extension to corridor-level analysis if computational resources permit.
- Development and testing of AI modules for perception, prediction, and control within the simulated environment.
- Comparative evaluation of AI-driven strategies against conventional traffic control methods using defined performance metrics.
- Consideration of mixed traffic conditions including cars, buses, matatus (public service vehicles), motorcycles, and pedestrians.

1.7.2 Limitations

The research acknowledges the following limitations:

- The simulation environment, while designed to be realistic, cannot fully capture all complexities of real-world traffic behavior and interactions.

- The study focuses on technical feasibility and performance evaluation, with limited consideration of institutional, political, and implementation barriers.
- Resource constraints may limit the scale and complexity of the simulated scenarios.
- The research does not include real-world deployment and validation, focusing instead on simulation-based proof of concept.

1.7.3 Delimitations

To maintain focus and feasibility, the study deliberately excludes:

- Large-scale city-wide traffic simulation due to computational constraints.
- Detailed economic analysis of implementation costs and benefits.
- Consideration of vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication systems, which are not yet deployed in Kenya.

1.8 Thesis Structure

This research proposal is organized into seven chapters. Chapter 1 has introduced the research background, problem statement, objectives, and significance. Chapter 2 presents a comprehensive review of relevant literature. Chapter 3 details the research methodology. Chapter 4 describes the proposed simulation framework design. Chapter 5 outlines the implementation approach. Chapter 6 discusses the evaluation strategy, and Chapter 7 presents the work plan and conclusion.

Chapter 2

Literature Review

2.1 Introduction to Urban Traffic Management

Urban traffic management represents a complex challenge at the intersection of civil engineering, computer science, urban planning, and social sciences. The fundamental goal of traffic management is to optimize the movement of people and goods through transportation networks while minimizing negative externalities such as congestion, pollution, and accidents (Zhang et al., 2018). Traditional approaches to traffic management have evolved from manual control to fixed-time signal systems, and more recently to adaptive systems that respond to real-time traffic conditions (Papageorgiou et al., 2003).

The complexity of urban traffic systems arises from several factors: the large number of interacting entities (vehicles, pedestrians, infrastructure), the dynamic nature of traffic flow, the influence of human behavior, and the impact of external factors such as weather events and special occasions (Helbing, 2001). This complexity necessitates sophisticated approaches to modeling, analysis, and control, which has led to the emergence of Intelligent Transportation Systems (ITS) as a multidisciplinary field dedicated to addressing urban mobility challenges through technological innovation (Ghaffarian et al., 2021).

2.2 Traffic Congestion in Developing Cities

Urban traffic congestion in developing cities presents distinct characteristics and challenges compared to developed contexts. Mohan and Tiwari (2017) identify several key factors that differentiate traffic conditions in cities like Nairobi, Mumbai, and Lagos from their developed counterparts:

Mixed Traffic Conditions: Developing cities typically feature heterogeneous traffic streams with significant variations in vehicle types, sizes, speeds, and operational characteristics. This includes the prominence of informal public transport systems, such as matatus in Kenya, which exhibit unique boarding/alighting patterns and routing behaviors (Kumar and Barrett, 2020).

Infrastructural Constraints: Many developing cities have transportation infrastructure that has not kept pace with rapid urbanization and population growth. This results in capacity constraints, inadequate road maintenance, and limited dedicated facilities for different transport modes (Otieno et al., 2021).

Institutional Challenges: Effective traffic management requires coordinated action across multiple government agencies, which can be challenging in contexts with limited institutional capacity, fragmented responsibilities, and resource constraints (Maitha and Nyong'o, 2020).

Behavioral Factors: Driver behavior in many developing cities is characterized by lower compliance with traffic regulations, more aggressive driving styles, and adaptation to informal traffic management practices (Salazar et al., 2021).

The economic impacts of congestion in developing cities are substantial. Kamau and Ochieng (2020) estimate that traffic congestion costs Nairobi approximately KES 100 billion annually in lost productivity, additional fuel consumption, and environmental damage. Similar figures have been reported for other African cities, including Lagos (Nigeria) and Accra (Ghana) (Bank, 2021).

2.3 Intelligent Transportation Systems (ITS)

Intelligent Transportation Systems represent an integrated approach to addressing transportation challenges through the application of advanced technologies. ITS encompasses a wide range of applications, including advanced traffic management systems, advanced traveler information systems, commercial vehicle operations, and advanced vehicle control systems (Ghaffarian et al., 2021).

The evolution of ITS can be traced through several generations (Zhang et al., 2018):

First Generation (1960s-1980s): Focused on isolated systems such as traffic signal controllers and basic surveillance systems. These systems operated with limited integration and minimal adaptive capabilities.

Second Generation (1980s-2000s): Saw the integration of multiple systems and the introduction of basic adaptive control strategies. This period also witnessed the emergence of traveler information systems and electronic toll collection.

Third Generation (2000s-Present): Characterized by the integration of communication technologies, data analytics, and artificial intelligence. Modern ITS leverage cloud computing, IoT devices, and machine learning to create increasingly sophisticated and responsive systems.

Key components of contemporary ITS include:

Sensing and Data Collection: Modern ITS rely on diverse data sources including inductive loops, video cameras, radar sensors, GPS trajectories, and crowd-sourced data from mobile devices (Guo et al., 2020).

Communication Infrastructure: Effective ITS require robust communication networks for transmitting data between vehicles, infrastructure, and control centers. Technologies include dedicated short-range communications (DSRC), cellular networks (4G/5G), and fiber optics (Abduljabbar et al., 2021).

Data Processing and Analytics: The volume and velocity of data generated by ITS necessitate advanced analytics capabilities, including real-time processing, machine learning, and predictive modeling (Lana et al., 2021).

Control and Actuation: The ultimate goal of ITS is to influence traffic flow through adaptive signal control, variable message signs, routing recommendations, and other interventions (Chu et al., 2019).

The implementation of ITS in developing contexts faces specific challenges, including high costs, technical complexity, maintenance requirements, and institutional barriers (Mohan and Tiwari, 2017). However, the potential benefits in terms of congestion reduction, safety improvement, and environmental sustainability make ITS an important area of research and development for cities in Kenya and similar contexts.

2.4 Simulation in Transportation Research

Simulation has emerged as a powerful tool in transportation research, enabling the study of complex systems under controlled conditions. Transportation simulations can be categorized into several types based on their scope and level of detail (Barceló, 2010):

Macroscopic Simulations: Focus on aggregate traffic flow characteristics, treating traffic as a continuous fluid. These models are computationally efficient but lack detail on individual vehicle behavior.

Microscopic Simulations: Model individual vehicles and their interactions, providing detailed insights into traffic dynamics. These models are more computationally intensive but offer greater realism.

Mesoscopic Simulations: Combine elements of both macroscopic and microscopic approaches, balancing computational efficiency with sufficient detail for many applications.

Submicroscopic Simulations: Include detailed vehicle dynamics and driver behavior models, often used for safety analysis and vehicle design.

Several simulation platforms have been developed specifically for transportation research. The Simulation of Urban Mobility (SUMO) is an open-source, microscopic, multi-modal traffic simulation package designed to handle large networks (Lopez et al., 2018). SUMO has been widely adopted in research due to its flexibility, extensibility, and active developer community. Other notable simulation tools include VISSIM (commercial software with detailed behavioral models), AIMSUN (commercial software with multi-resolution capabilities), and MATSim (open-source agent-based simulation framework) (Barceló, 2010).

The application of simulation in transportation research spans multiple domains:

Traffic Signal Optimization: Simulation enables the testing of signal timing plans under various traffic conditions without disrupting real-world operations (Gao et al., 2020).

Network Planning: Simulations can assess the impact of new infrastructure, road modifications, or policy changes on network performance (Zhang et al., 2019).

Emergency Evacuation Planning: Simulations help design effective evacuation strategies for natural disasters or other emergencies (Liu et al., 2020).

Environmental Impact Assessment: Traffic simulations can estimate emissions and fuel consumption under different scenarios (Cooper and Barlow, 2021).

In the context of developing cities, simulation faces specific challenges related to data availability, model calibration, and the representation of unique local conditions (Salazar et al., 2021). However, when properly implemented, simulation offers a valuable approach for developing and testing traffic management strategies tailored to these contexts.

2.5 Artificial Intelligence in Traffic Management

Artificial intelligence has transformed numerous domains, and traffic management is no exception. AI techniques applied to traffic management can be categorized into several areas:

2.5.1 Computer Vision for Traffic Perception

Computer vision techniques enable the extraction of traffic information from video feeds and other visual data sources. Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in tasks such as vehicle detection, classification, and tracking (Li et al., 2016). Popular architectures include YOLO (You Only Look Once), SSD (Single Shot Detector), and Faster R-CNN (Region-Based Convolutional Neural Network) (Redmon and Farhadi, 2018).

The application of computer vision to traffic management faces challenges including varying lighting conditions, occlusions, and the need for real-time processing (Guo et al., 2020). However, advances in deep learning and hardware acceleration have made vision-based traffic analysis increasingly feasible for practical applications.

2.5.2 Predictive Modeling for Traffic Flow

Predictive models anticipate future traffic conditions based on historical patterns and real-time data. Traditional time-series approaches such as ARIMA (AutoRegressive Integrated Moving Average) have been supplemented and often outperformed by machine learning methods including Support Vector Machines (SVM), Random Forests, and neural networks (Vlahogianni et al., 2014).

More recently, deep learning architectures such as Long Short-Term Memory (LSTM) networks and Transformers have shown exceptional performance in capturing complex temporal dependencies in traffic data (Zhao et al., 2020). These models can incorporate multiple data sources and external factors to improve prediction accuracy.

2.5.3 Reinforcement Learning for Traffic Control

Reinforcement learning (RL) has emerged as a promising approach for adaptive traffic control. In RL frameworks, an agent learns optimal control policies through interaction with the environment, receiving rewards for desirable outcomes such as reduced delays or increased throughput (Chu et al., 2019).

RL applications in traffic control range from isolated intersection control to network-wide optimization. Challenges include the large state and action spaces, the need for efficient exploration strategies, and ensuring stability and safety (Wei et al., 2019). Despite these challenges, RL-based approaches have demonstrated significant improvements over conventional control methods in simulation studies (Gao et al., 2020).

2.5.4 Multi-Agent Systems for Distributed Control

Multi-agent systems (MAS) provide a framework for distributed traffic control, where multiple intelligent agents coordinate to optimize network performance. MAS approaches can be more scalable and robust than centralized control systems, particularly for large networks (Abduljabbar et al., 2021).

Challenges in MAS for traffic control include designing effective communication protocols, ensuring coordination among agents, and avoiding destabilizing interactions (Wei et al., 2019). However, the distributed nature of MAS makes them well-suited for the increasingly connected and automated transportation systems of the future.

2.6 Digital Twins in Transportation

Digital twins represent virtual replicas of physical systems that are continuously updated with real-time data, enabling simulation, analysis, and control (Cheng et al., 2020). In transportation, digital twins integrate multiple data sources, simulation models, and analytical tools to create comprehensive virtual representations of transportation systems.

The application of digital twins in transportation offers several advantages (Kaplan and Haenlein, 2021):

Real-time Monitoring and Diagnosis: Digital twins can continuously compare simulated and actual system behavior to detect anomalies and diagnose problems.

Predictive Analytics: By running simulations into the future, digital twins can anticipate congestion, identify potential bottlenecks, and evaluate mitigation strategies.

Scenario Testing: Digital twins provide a safe environment for testing new control strategies, infrastructure modifications, or policy changes before implementation.

Decision Support: The insights generated by digital twins can inform operational decisions and long-term planning.

The development of transportation digital twins faces challenges including data integration, model fidelity, computational requirements, and the need for specialized expertise (Ghaffarian et al., 2021). However, as data availability and computational capabilities continue to improve, digital twins are likely to play an increasingly important role in transportation management.

In the context of developing cities, digital twins offer particular promise for overcoming data limitations through the integration of multiple data sources and the use of synthetic data generation techniques (Salazar et al., 2021). This research contributes to this emerging field by developing a digital twin framework specifically designed for Kenyan urban environments.

2.7 Traffic Management in the Kenyan Context

Research on traffic management in Kenya has primarily focused on descriptive analyses of congestion patterns, economic impacts, and policy recommendations. Otieno et al. (2021) provide a comprehensive assessment of traffic congestion in Nairobi, identifying key congestion hotspots, peak periods, and contributing factors. Their research highlights the significant economic costs of congestion but offers limited technical solutions.

Maitha and Nyong'o (2020) examine the institutional framework for traffic management in Kenya, identifying fragmentation among multiple agencies and limited technical capacity as major constraints. They recommend institutional reforms and increased investment in traffic management infrastructure but do not explore technological innovations in detail.

Several studies have documented the unique characteristics of Kenyan traffic, particularly the role of matatus (informal public transport vehicles) in shaping traffic dynamics (Kumar and Barrett, 2020). These studies highlight the need for traffic management approaches that account for the specific behaviors and operational patterns of matatus, which differ significantly from formal public transport systems in other contexts.

Despite this growing body of research, there remains a significant gap in studies exploring the application of advanced technologies, particularly AI and simulation, to traffic management in Kenya. This research aims to address this gap by developing a comprehensive framework for evaluating AI-powered traffic control strategies in simulated Kenyan urban environments.

2.8 Research Gaps and Contribution

Based on the literature review, several research gaps have been identified:

1. **Context-Specific Solutions:** Most AI-powered traffic management research has focused on developed country contexts with different traffic patterns, infrastructure, and behavioral norms. There is limited research on adapting these approaches to developing cities like those in Kenya.

2. **Data Scarcity Solutions:** While data scarcity is a recognized challenge in developing contexts, there is limited research on systematic approaches for overcoming this barrier, particularly through synthetic data generation and simulation.

3. **Integrated Evaluation Frameworks:** Existing studies often focus on isolated aspects of traffic

management (e.g., signal control or prediction) without integrated frameworks that address the full pipeline from perception to control.

4. **Practical Implementation Pathways:** There is a gap between theoretical research on AI for traffic management and practical implementation strategies suited to the institutional and resource constraints of developing cities.

This research contributes to addressing these gaps by:

1. Developing a simulation framework specifically designed for Kenyan urban traffic conditions.

2. Creating integrated AI modules for perception, prediction, and control within this framework.

3. Establishing a comprehensive evaluation methodology that assesses multiple dimensions of performance.

4. Providing insights into the practical implementation of simulation-validated strategies in real-world Kenyan contexts.

Chapter 3

Theoretical Framework

3.1 Introduction

This chapter presents the theoretical foundations underpinning the proposed research. The framework integrates concepts from multiple disciplines including transportation engineering, computer science, complex systems theory, and urban planning. By establishing this theoretical foundation, the research aims to ensure that the proposed simulation framework is grounded in established principles while incorporating innovative approaches suited to the Kenyan context.

3.2 Complex Systems Theory in Transportation

Urban transportation networks exemplify complex systems characterized by emergent behavior, non-linear dynamics, and adaptation (Helbing, 2001). Complex systems theory provides a valuable lens for understanding traffic flow phenomena that cannot be adequately explained through reductionist approaches.

Key principles from complex systems theory relevant to this research include:

Emergence: System-level properties (e.g., congestion patterns) emerge from the interactions of individual components (vehicles, drivers) following relatively simple rules. This bottom-up perspective informs the agent-based modeling approach used in the simulation framework.

Non-linearity: Small changes in system parameters can lead to disproportionate effects due to feedback loops and threshold effects. This necessitates careful calibration of simulation models and sensitivity analysis.

Adaptation: System components adapt their behavior based on experience and environmental cues. This principle underpins the learning mechanisms in both the simulated driver behavior models and the AI control algorithms.

The application of complex systems theory to transportation has led to important insights such as the three-phase traffic theory (Kerner, 2004) and the understanding of capacity drop phenomena

(Helbing, 2001). These theoretical concepts inform the design of the simulation scenarios and the interpretation of results.

3.3 Queueing Theory and Traffic Flow Modeling

Queueing theory provides mathematical models for analyzing waiting lines and congestion in systems with limited capacity. In transportation applications, queueing theory helps model intersection delays, bottleneck formation, and network performance (May, 1990).

Fundamental queueing theory concepts relevant to this research include:

Little's Law: This fundamental relationship (L = W) states that the average number of customers in a stable system (L) equals the arrival rate () multiplied by the average time each customer spends in the system (W). This principle provides a basis for estimating delays from queue lengths and vice versa.

Queueing Models: Different queueing models (M/M/1, M/D/1, etc.) offer approximations for various traffic conditions. While real-world traffic often deviates from the assumptions of classical queueing models, these models provide valuable benchmarks and analytical tools.

Network Queueing Theory: Extensions of queueing theory to networks help analyze systems with multiple interconnected queues, such as urban road networks with signalized intersections.

The proposed research will leverage queueing theory principles in designing the evaluation metrics and interpreting simulation results, particularly for comparing the performance of different control strategies under varying demand levels.

3.4 Machine Learning Foundations

The AI components of the proposed framework are grounded in machine learning theory. Key theoretical concepts include:

3.4.1 Statistical Learning Theory

Statistical learning theory provides a framework for understanding the generalization performance of machine learning algorithms (Vapnik, 1999). Concepts such as the bias-variance tradeoff, VC dimension, and structural risk minimization inform the design and evaluation of the AI models used in this research.

The challenge of limited real-world data in the Kenyan context makes theoretical understanding of generalization particularly important. Techniques such as regularization, cross-validation, and ensemble methods will be employed to enhance model robustness despite data constraints.

3.4.2 Deep Learning Theory

Deep learning models, particularly convolutional neural networks (CNNs) for perception tasks and recurrent neural networks (RNNs) for prediction tasks, are central to the proposed framework. Theoretical understanding of concepts such as gradient-based optimization, backpropagation, and architectural design choices informs the implementation of these components (Goodfellow et al., 2016).

Recent advances in deep learning theory, including attention mechanisms, transformer architectures, and self-supervised learning, offer opportunities for improving model performance while reducing data requirements (Vaswani et al., 2017). These advances will be explored where applicable to the specific challenges of Kenyan traffic data.

3.4.3 Reinforcement Learning Theory

Reinforcement learning (RL) provides the theoretical foundation for the adaptive control component of the framework. Key RL concepts include:

Markov Decision Processes (MDPs): MDPs formalize sequential decision-making problems with states, actions, transitions, and rewards. Traffic signal control can be framed as an MDP where states represent traffic conditions, actions represent signal phase selections, and rewards reflect performance metrics (Sutton and Barto, 2018).

Value Functions and Policy Optimization: RL algorithms learn either value functions (estimating expected cumulative reward) or policies (mapping states to actions) directly. The tradeoffs between value-based, policy-based, and actor-critic approaches will be evaluated for the traffic control application.

Exploration-Exploitation Tradeoff: RL agents must balance exploring new actions to discover better strategies with exploiting known good actions. This tradeoff is particularly important in traffic control where poor decisions can cause significant disruptions.

The theoretical understanding of RL convergence properties, sample efficiency, and safety considerations will guide the design and implementation of the adaptive control module.

3.5 Simulation Theory and Methodology

Simulation methodology provides the theoretical foundation for creating and validating digital representations of real-world systems. Key concepts include:

Model Verification and Validation: Verification ensures that the simulation implementation correctly represents the conceptual model, while validation ensures that the model adequately represents the real-world system (Law, 2015). These processes are essential for establishing the credibility of simulation results.

Experimental Design: Proper experimental design, including factor selection, replication strategies, and statistical analysis methods, ensures that simulation experiments yield meaningful and generalizable results (Kleijnen, 2015).

Output Analysis: Simulation output analysis techniques, including confidence interval estimation, variance reduction, and sensitivity analysis, help extract robust insights from stochastic simulation results (Law, 2015).

The proposed research will adhere to established simulation methodology principles to ensure the reliability and validity of the findings.

3.6 Integration Framework

The theoretical framework integrates these diverse theoretical perspectives through a systems approach that recognizes the multi-level nature of urban transportation systems. The integration is visualized in Figure 3.1.

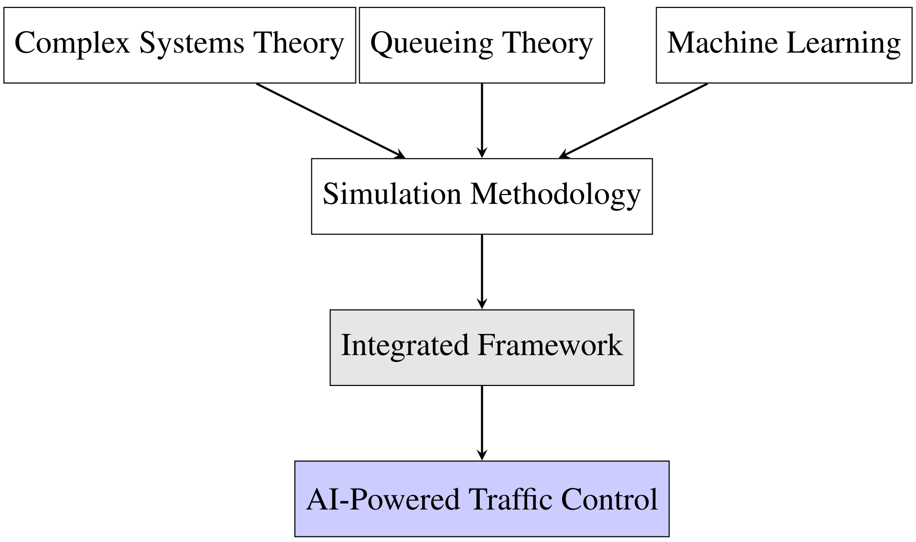

Figure 3.1: Theoretical Integration Framework

This integrated theoretical framework ensures that the research is grounded in established principles while innovatively addressing the specific challenges of Kenyan urban traffic management.

Chapter 4

Research Methodology

4.1 Research Design

This research adopts a design science research methodology, which focuses on the creation and evaluation of artifacts designed to solve identified problems (Hevner et al., 2004). The primary artifact is the simulation framework for evaluating AI-powered traffic control strategies. The research follows a cyclical process of design, development, evaluation, and refinement, as illustrated in Figure 4.1.

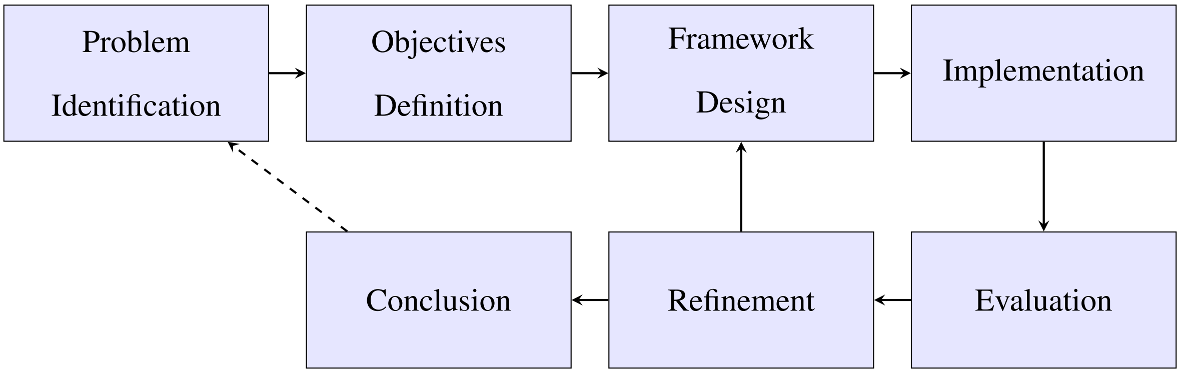

Figure 4.1: Research Design Process

The research incorporates both quantitative and qualitative elements within a primarily simulation-based experimental approach. Quantitative methods dominate the evaluation phase, while qualitative approaches inform the design and contextual interpretation of results.

4.2 Simulation Environment Setup

The simulation environment will be developed using the Simulation of Urban Mobility (SUMO), an open-source, microscopic, multi-modal traffic simulation package (Lopez et al., 2018). SUMO provides the necessary flexibility and extensibility for implementing custom control strategies and integrating with external AI modules.

4.2.1 Network Modeling

A digital twin of a representative Kenyan intersection will be created based on geometric and operational characteristics of actual intersections in Nairobi or Nakuru. The modeling process includes:

Network Definition: Creating the road network topology including lane configurations, curvature, gradients, and intersection geometry.

Traffic Demand Modeling: Defining origin-destination matrices based on traffic flow patterns observed in Kenyan urban environments.

Vehicle Type Definitions: Creating accurate representations of the vehicle mix in Kenyan traffic, including:

- Personal cars (sedans, SUVs)
- Matatus (14-seater and 25-seater minibuses)
- Motorcycles (boda-bodas)
- Buses (public and private)
- Trucks (light and heavy)

Driver Behavior Modeling: Calibrating driver behavior parameters (car-following, lane-changing, gap acceptance) to reflect Kenyan driving patterns based on available literature and observational data.

4.2.2 Simulation Scenarios

Multiple simulation scenarios will be developed to evaluate the framework under different conditions:

Baseline Scenario: Represents current traffic conditions with conventional signal control.

Peak Hour Scenarios: High-demand conditions during morning and evening peak periods.

Special Event Scenarios: Conditions with unusual demand patterns (e.g., holidays, events).

Incident Scenarios: Conditions with reduced capacity due to accidents or roadworks.

4.3 Data Generation and Management

Addressing data scarcity is a core objective of this research. The data generation approach includes:

4.3.1 Synthetic Data Generation

Algorithm 1 outlines the synthetic data generation process.

Algorithm 1 Synthetic Traffic Data Generation

Initialize simulation environment with network configuration
for each vehicle type in vehicle mix **do**
    Define vehicle characteristics (length, width, acceleration, deceleration)
    Define behavior parameters for vehicle type
end for
for each simulation scenario **do**
    Set traffic demand parameters
    Run simulation with conventional control
    Extract vehicle trajectories, speeds, accelerations
    Extract signal state information
    Extract queue lengths and waiting times
    Annotate data with ground truth labels
end for
Aggregate data across scenarios
Partition data into training, validation, and test sets
Apply data augmentation techniques (rotation, scaling, noise addition)

The synthetic data will include:

- Vehicle trajectories (position, speed, acceleration over time)
- Traffic state variables (density, flow, occupancy)
- Signal timing and phase information
- Performance metrics (delay, stops, queue lengths)

4.3.2 Data Quality Assurance

Data quality will be ensured through:

- Validation against available real-world data where possible

- Sensitivity analysis to assess the impact of parameter variations
- Cross-validation techniques to evaluate model robustness

4.4 AI Module Development

The framework includes three core AI modules: perception, prediction, and control.

4.4.1 Perception Module

The perception module is responsible for detecting and tracking vehicles within the simulation environment. The module will be implemented using deep learning approaches, particularly convolutional neural networks (CNNs) for object detection.

The implementation will follow the architecture shown in Figure 4.2.

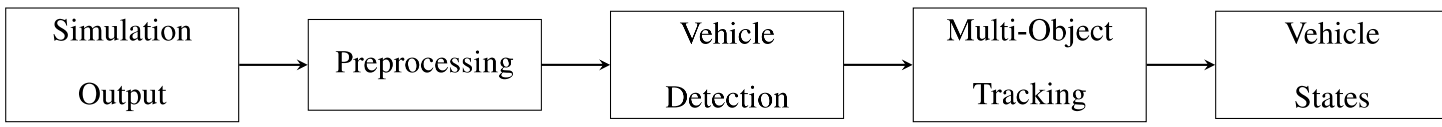

Figure 4.2: Perception Module Architecture

4.4.2 Prediction Module

The prediction module forecasts short-term traffic conditions based on historical patterns and current state. The module will utilize time-series forecasting techniques, with Long Short-Term Memory (LSTM) networks as the primary approach.

The prediction process is formalized in Algorithm 2.

Algorithm 2 Traffic Flow Prediction using LSTM

```
Initialize LSTM model architecture
Load historical traffic data
Preprocess data (normalization, sequence formatting)
Train LSTM model on historical data
Validate model performance on test set
while simulation is running do
    Get current traffic state from perception module
    Format input sequence for prediction
    Generate predictions for next time steps (5-15 minutes)
    Update prediction confidence intervals
end while
```

4.4.3 Control Module

The control module implements adaptive signal control using reinforcement learning. The module will be based on the Deep Q-Network (DQN) algorithm, which combines Q-learning with deep neural networks for value function approximation.

The control module operation is detailed in Algorithm 3.

Algorithm 3 Reinforcement Learning for Adaptive Traffic Control

Initialize DQN with random weights
Initialize replay memory
for each episode **do**
  Reset simulation to initial state
  for each time step **do**
    Observe current traffic state s_t
    Select action a_t using -greedy policy
    Execute action in simulation environment
    Observe next state s_{t+1} and reward r_t
    Store transition (s_t, a_t, r_t, s_{t+1}) in replay memory
    Sample random minibatch from replay memory
    Perform gradient descent step on DQN
    Update target network periodically
  end for
end for

4.5 Integration Framework

The AI modules will be integrated into a cohesive framework using the TraCI (Traffic Control Interface) API provided by SUMO. This interface allows external applications to control the simulation and access simulation data in real-time.

The integration architecture is shown in Figure 4.3.

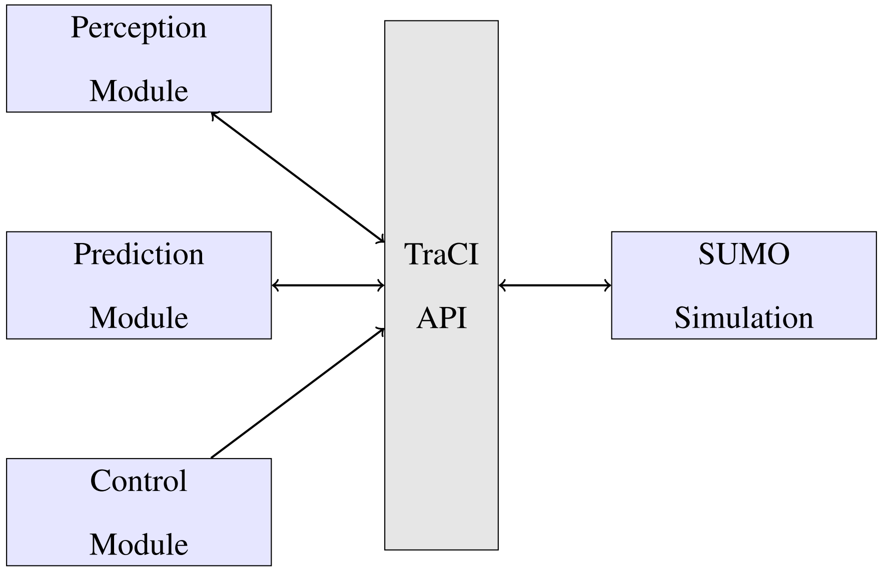

Figure 4.3: Framework Integration Architecture

4.6 Evaluation Methodology

The evaluation will compare the performance of the AI-powered control strategies against conventional methods across multiple metrics.

4.6.1 Performance Metrics

The following metrics will be used for evaluation:

Efficiency Metrics:

- Average delay per vehicle
- Total travel time of stops
- Throughput (vehicles per hour)

Environmental Metrics:

- Fuel consumption
- CO2 emissions
- Other pollutant emissions (NOx, PM)

Safety Metrics:

- Conflict points

- Deceleration rates -to-collision indicators

Equity Metrics:

- Delay distribution across vehicle types
- Impact on public transport priority

4.6.2 Statistical Analysis

The evaluation will include statistical analysis to ensure the robustness of results:

Descriptive Statistics: Summary statistics for all performance metrics under different control strategies.

Inferential Statistics: Hypothesis testing (t-tests, ANOVA) to determine statistically significant differences between control strategies.

Sensitivity Analysis: Assessment of how results vary with changes in key parameters (demand levels, behavior parameters).

4.7 Ethical Considerations

This research adheres to ethical principles in research involving simulation and AI:

Data Privacy: The use of synthetic data eliminates privacy concerns associated with real-world traffic data.

Algorithmic Fairness: The evaluation includes equity metrics to ensure that control strategies do not disproportionately disadvantage specific user groups.

Transparency: The open-source nature of the framework promotes transparency and reproducibility.

Responsible AI: The research considers the societal implications of AI-powered traffic control and includes safeguards against potential negative consequences.

4.8 Limitations of the Methodology

The methodology has several limitations:

Simulation Reality Gap: Despite efforts to create realistic models, simulations inevitably simplify real-world complexity.

Computational Constraints: The complexity of AI models and large-scale simulations may require computational resources that limit scenario complexity.

Generalizability: Results from a single intersection simulation may not fully generalize to network-wide applications.

These limitations will be explicitly acknowledged in the interpretation of results and recommendations for future research.

Chapter 5

Implementation Plan

5.1 Technical Architecture

The implementation will follow a modular architecture that separates concerns and enables independent development of system components. The overall architecture is depicted in Figure 5.1.

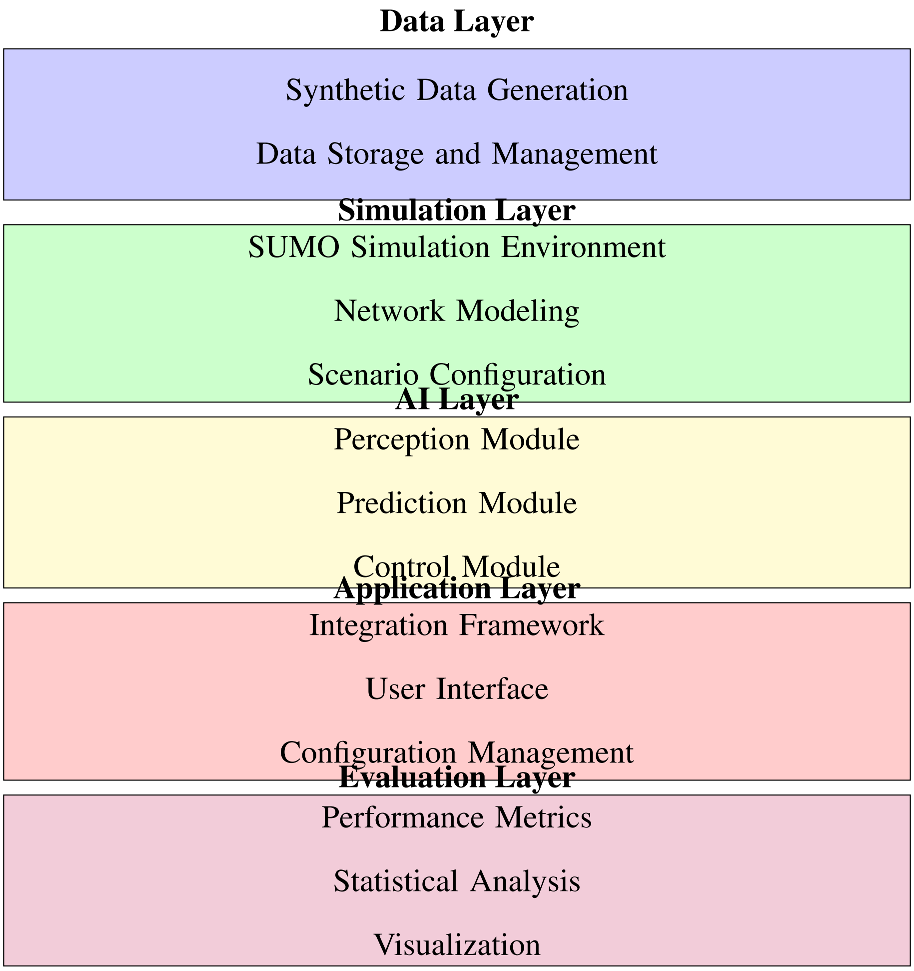

Figure 5.1: System Architecture Overview

5.2 Development Environment

The development will utilize the following tools and technologies:

Programming Languages:

- Python 3.8+ (primary language for AI modules and integration)
- XML (for SUMO configuration files)

Libraries and Frameworks:

- TensorFlow/PyTorch for deep learning implementations
- OpenCV for computer vision tasks
- NumPy, Pandas for data manipulation
- Matplotlib, Seaborn for visualization
- Scikit-learn for traditional ML algorithms

Simulation Tools:

- SUMO (Simulation of Urban Mobility)
- TraCI (Traffic Control Interface) for external control

Development Tools:

- Git for version control
- Docker for environment consistency
- Jupyter Notebooks for exploratory analysis

5.3 Implementation Timeline

The implementation will follow the timeline outlined in Table 5.1.

Phase	**Activities and Deliverables**
Months 1-2	**Environment Setup and Literature Review** • Install and configure development environment • Conduct comprehensive literature review • Define detailed technical specifications

	• Deliverable: Annotated bibliography and technical specification document
Months 3-4	**Simulation Environment Development** • Create digital twin of representative intersection • Model vehicle types and behavior parameters • Develop simulation scenarios • Deliverable: Functional simulation environment with baseline scenarios
Months 5-6	**Data Generation and Perception Module** • Implement synthetic data generation pipeline • Develop and train vehicle detection models • Implement multi-object tracking • Deliverable: Synthetic dataset and trained perception models
Months 7-8	**Prediction and Control Modules** • Implement traffic flow prediction using LSTM • Develop reinforcement learning control agent • Integrate modules with simulation environment • Deliverable: Functional AI modules with initial integration
Months 9-10	**System Integration and Testing** • Complete framework integration • Conduct systematic testing and debugging • Optimize performance and resource usage • Deliverable: Fully integrated simulation framework
Months 11-12	**Evaluation and Documentation**

	• Execute comprehensive evaluation experiments • Analyze results and draw conclusions • Prepare final documentation and thesis • Deliverable: Complete thesis and research publications

Table 5.1: Implementation Timeline

5.4 Code Structure and Organization

The codebase will be organized following software engineering best practices for maintainability and reproducibility. The directory structure will be as follows:

```
traffic_ai_framework/
          data/                        # Data storage
                raw/                    # Raw simulation outputs
                processed/              # Processed datasets
                models/                 # Trained model files
          src/                        # Source code
                simulation/             # Simulation configuration
                perception/             # Vehicle detection and tracking
                prediction/             # Traffic flow forecasting
                control/               # Adaptive signal control
                integration/           # Framework integration
                evaluation/            # Performance evaluation
          tests/                      # Unit and integration tests
          docs/                       # Documentation
          config/                     # Configuration files
          scripts/                    # Utility scripts
```

Listing 5.1: Project Directory Structure

5.5 Key Implementation Details

5.5.1 SUMO Integration

The integration with SUMO will use the TraCI API as shown in the following code snippet:

```
import traci
```

```python
import sumolib

class SUMOController:
    def __init__(self, config_file):
        self.sumo_cmd = ["sumo", "-c", config_file]
        self.connection = None

    def start_simulation(self):
        """Start the SUMO simulation"""
        traci.start(self.sumo_cmd)
        self.connection = traci

    def get_vehicle_data(self):
        """Retrieve vehicle data from simulation"""
        vehicle_ids = self.connection.vehicle.getIDList()
        vehicle_data = {}
        for veh_id in vehicle_ids:
            position = self.connection.vehicle.getPosition(veh_id)
            speed = self.connection.vehicle.getSpeed(veh_id)
            vehicle_data[veh_id] = {
                'position': position,
                'speed': speed,
                'type': self.connection.vehicle.getTypeID(veh_id)
            }
        return vehicle_data

    def set_traffic_light_phase(self, tl_id, phase):
        """Set traffic light phase"""
        self.connection.trafficlight.setPhase(tl_id, phase)

    def step_simulation(self):
        """Advance simulation by one step"""
        self.connection.simulationStep()
```

Listing 5.2: SUMO Integration Example

5.5.2 Perception Module Implementation

The perception module will use a YOLO-based approach for vehicle detection:

```python
import cv2
import numpy as np
import tensorflow as tf

class VehicleDetector:
    def __init__(self, model_path):
        self.model = tf.keras.models.load_model(model_path)
        self.classes = ['car', 'truck', 'motorcycle', 'bus']

    def preprocess_image(self, image):
        """Preprocess image for YOLO model"""
        image = cv2.resize(image, (416, 416))
        image = image / 255.0
        image = np.expand_dims(image, axis=0)
        return image

    def detect_vehicles(self, image):
        """Detect vehicles in image"""
        processed_image = self.preprocess_image(image)
        predictions = self.model.predict(processed_image)
        detections = self.postprocess_predictions(predictions)
        return detections

    def postprocess_predictions(self, predictions):
        """Convert model outputs to bounding boxes"""
        # Implementation of YOLO postprocessing
        pass
```

Listing 5.3: Perception Module Implementation

5.5.3 Reinforcement Learning Controller

The RL controller will implement a DQN algorithm:

```python
import numpy as np
import tensorflow as tf
from collections import deque
import random

```

```python
class DQNTrafficController:
    def __init__(self, state_size, action_size):
        self.state_size = state_size
        self.action_size = action_size
        self.memory = deque(maxlen=2000)
        self.gamma = 0.95    # discount rate
        self.epsilon = 1.0   # exploration rate
        self.epsilon_min = 0.01
        self.epsilon_decay = 0.995
        self.learning_rate = 0.001
        self.model = self._build_model()
        self.target_model = self._build_model()

    def _build_model(self):
        """Build neural network for Q-value approximation"""
        model = tf.keras.Sequential()
        model.add(tf.keras.layers.Dense(24, input_dim=self.state_size, activation='relu'))
        model.add(tf.keras.layers.Dense(24, activation='relu'))
        model.add(tf.keras.layers.Dense(self.action_size, activation='linear'))
        model.compile(loss='mse', optimizer=tf.keras.optimizers.Adam(lr=self.learning_rate))
        return model

    def remember(self, state, action, reward, next_state, done):
        """Store experience in replay memory"""
        self.memory.append((state, action, reward, next_state, done))

    def act(self, state):
        """Select action using epsilon-greedy policy"""
        if np.random.rand() <= self.epsilon:
            return random.randrange(self.action_size)
        act_values = self.model.predict(state)
        return np.argmax(act_values[0])

    def replay(self, batch_size):
        """Train model on random batch from memory"""
        minibatch = random.sample(self.memory, batch_size)
```

```python
        for state, action, reward, next_state, done in minibatch:
            target = reward
            if not done:
                target = reward + self.gamma * np.amax(self.target_model.predict(next_state)[0])
            target_f = self.model.predict(state)
            target_f[0][action] = target
            self.model.fit(state, target_f, epochs=1, verbose=0)
        if self.epsilon > self.epsilon_min:
            self.epsilon *= self.epsilon_decay

    def update_target_model(self):
        """Update target network with primary network weights"""
        self.target_model.set_weights(self.model.get_weights())
```

Listing 5.4: Reinforcement Learning Controller

5.6 Testing Strategy

A comprehensive testing strategy will ensure the reliability and validity of the implementation:

Unit Testing: Individual components will be tested in isolation using frameworks like pytest.

Integration Testing: The interaction between modules will be tested to ensure proper integration.

System Testing: The complete framework will be tested end-to-end with various scenarios.

Performance Testing: Computational efficiency and resource usage will be evaluated.

Example unit test for the perception module:

```python
import unittest
from src.perception.vehicle_detector import VehicleDetector

class TestVehicleDetector(unittest.TestCase):
    def setUp(self):
        self.detector = VehicleDetector('test_model.h5')
        self.test_image = np.random.rand(640, 480, 3)

    def test_detection_output_format(self):
        detections = self.detector.detect_vehicles(self.test_image)
        self.assertIsInstance(detections, list)
        for detection in detections:
```

```python
            self.assertIn('class', detection)
            self.assertIn('confidence', detection)
            self.assertIn('bbox', detection)

    def test_preprocessing(self):
        processed = self.detector.preprocess_image(self.test_image)
        self.assertEqual(processed.shape, (1, 416, 416, 3))

if __name__ == '__main__':
    unittest.main()
```

Listing 5.5: Example Unit Test

5.7 Documentation and Reproducibility

Comprehensive documentation will ensure the reproducibility and usability of the framework:

Code Documentation: Inline comments and docstrings following PEP 257 conventions.

API Documentation: Automatically generated documentation using Sphinx.

User Guide: Step-by-step instructions for setting up and using the framework.

Research Compendium: Complete documentation of experiments, parameters, and results.

The implementation will adhere to FAIR principles (Findable, Accessible, Interoperable, Reusable) to maximize the impact and utility of the research outputs.

Chapter 6

Expected Results and Evaluation

6.1 Expected Outcomes

This research is expected to yield several significant outcomes that contribute to both academic knowledge and practical applications in traffic management.

6.1.1 Technical Outcomes

1. **Open-Source Simulation Framework:** A fully functional, open-source framework for evaluating AI-powered traffic control strategies in Kenyan urban environments. The framework will be publicly available to facilitate further research and development.

2. **Synthetic Kenyan Traffic Dataset:** A comprehensive synthetic dataset representing traffic conditions in Kenyan cities, addressing the current data scarcity challenge. The dataset will include vehicle trajectories, traffic states, and performance metrics under various scenarios.

3. **Validated AI Models:** Trained and validated AI models for traffic perception, prediction, and control that are specifically adapted to Kenyan traffic conditions. The models will demonstrate state-of-the-art performance on the defined evaluation metrics.

4. **Performance Benchmarks:** Comparative analysis of AI-powered control strategies against conventional methods, providing evidence-based insights into their relative effectiveness in Kenyan contexts.

6.1.2 Academic Outcomes

1. **Master's Thesis:** A comprehensive document detailing the research methodology, implementation, results, and conclusions, submitted in fulfillment of the Master of Science degree requirements.

2. **Research Publications:** At least two peer-reviewed journal articles or conference papers disseminating the research findings to the academic community.

3. **Contribution to Theory:** Theoretical insights into the application of AI techniques in developing urban contexts and the validity of simulation-based approaches for traffic management research.

6.1.3 Practical Outcomes

1. **Implementation Guidelines:** Evidence-based guidelines for the potential implementation of AI-powered traffic control systems in Kenyan cities, including technical requirements, cost estimates, and institutional considerations.

2. **Policy Recommendations:** Recommendations for policymakers and traffic authorities on integrating simulation-based approaches into traffic management planning and decision-making processes.

3. **Capacity Building:** Development of local expertise in AI and simulation for transportation applications, contributing to Kenya's technological capacity in this domain.

6.2 Evaluation Framework

The evaluation framework will systematically assess the performance of the proposed AI-powered traffic control strategies against conventional methods across multiple dimensions.

6.2.1 Performance Metrics

The evaluation will utilize the comprehensive set of metrics defined in Section 4.6, organized into four categories as shown in Table 6.1.

Category	Metric	Description
Efficiency	Average Delay per Vehicle	Total delay divided by number of vehicles (seconds)
	Travel Time	Mean travel time through intersection (seconds)

	Number of Stops	Average number of stops per vehicle
	Throughput	Vehicles passing through intersection per hour
	Queue Length	Maximum and average queue lengths (meters)
Environmental	Fuel Consumption	Total fuel consumed (liters)
	CO2 Emissions	Total carbon dioxide emissions (kg)
	Other Emissions	Nitrogen oxides (NOx) and particulate matter (PM)
	Noise Pollution	Estimated noise levels based on traffic flow
Safety	Time-to-Collision	Minimum time to potential collision (seconds)
	Deceleration Rates	Frequency of harsh braking events
	Conflict Points	Number of potential vehicle conflicts
	Safety Margin	Average distance between vehicles
Equity	Delay Distribution	Variation in delays across vehicle types
	Public Transport Priority	Impact on public transport vehicles
	Pedestrian Wait Times	Average waiting time for pedestrians
	Equity Index	Gini coefficient of delay distribution

Table 6.1: Performance Metrics for Evaluation

6.2.2 Comparative Analysis

The evaluation will compare the following control strategies:

1. **Fixed-Time Control:** Conventional pre-timed signal control based on historical traffic patterns.
2. **Actuated Control:** Traffic-responsive control using induction loops or other detectors.
3. **AI-Powered Adaptive Control:** The proposed reinforcement learning-based control strategy.
4. **Hybrid Approaches:** Combinations of conventional and AI-based methods.

The comparison will be conducted under various traffic conditions (low, medium, high demand) and scenarios (normal, incident, special event).

6.2.3 Statistical Evaluation

The statistical evaluation will include:

Descriptive Statistics: Mean, median, standard deviation, and range for all performance metrics under each control strategy.

Inferential Statistics: Hypothesis testing to determine statistically significant differences between control strategies. For example, paired t-tests will compare the mean delays under AI control versus fixed-time control:

$$H_0 : \mu_{\text{AI}} = \mu_{\text{Fixed}} \quad \text{vs.} \quad H_1 : \mu_{\text{AI}} < \mu_{\text{Fixed}} \tag{6.1}$$

Effect Size Analysis: Calculation of effect sizes (e.g., Cohen's d) to quantify the magnitude of differences between strategies.

Confidence Intervals: 95% confidence intervals for key performance metrics to indicate estimation precision.

6.3 Expected Performance Results

Based on prior research in similar contexts (Chu et al., 2019; Gao et al., 2020), the AI-powered control strategy is expected to demonstrate significant improvements over conventional methods. The specific expected results are summarized in Table 6.3.

Performance Metric	**Fixed-Time Control**	**Actuated Control**	**AI-Powered Control (Expected)**

Average Delay per Vehicle (seconds)	45-60	35-50	25-35 (20-40% improvement)
Throughput (vehicles/hour)	800-1000	900-1100	1000-1200 (10-20% improvement)
Fuel Consumption (liters/hour)	40-50	35-45	30-38 (15-25% reduction)
CO_2 Emissions (kg/hour)	100-125	85-110	75-95 (15-25% reduction)
Number of Stops (per vehicle)	1.5-2.0	1.2-1.7	0.8-1.3 (25-35% reduction)
Queue Length (meters)	80-120	60-100	40-80 (30-40% reduction)

6.4 Sensitivity Analysis

Sensitivity analysis will assess the robustness of the results to variations in key parameters:

Traffic Demand: How performance changes with different demand levels (under-saturated, saturated, over-saturated conditions).

Vehicle Mix: Impact of variations in the proportion of different vehicle types (cars, matatus, motorcycles).

Behavior Parameters: Sensitivity to changes in driver behavior parameters (aggressiveness, compliance).

Incident Scenarios: Performance under non-recurrent congestion conditions.

The sensitivity analysis will help identify the conditions under which AI-powered control is most effective and any potential limitations or edge cases.

6.5 Validation Approach

While the primary validation will be within the simulation environment, several approaches will be used to enhance the validity of the results:

Model Calibration: The simulation model will be calibrated against available real-world data to ensure realistic representation of traffic dynamics.

Expert Validation: Domain experts will review the simulation scenarios, parameter settings, and results interpretation.

Cross-Validation: The AI models will be evaluated using cross-validation techniques to assess generalization performance.

Comparison with Literature: Results will be compared with findings from similar studies in other contexts to assess consistency and contextual differences.

6.6 Limitations and Mitigations

The evaluation has several limitations that will be acknowledged and addressed:

Simulation Reality Gap: The simulated environment may not fully capture all aspects of real-world traffic behavior. This will be mitigated through careful model calibration and sensitivity analysis.

Computational Constraints: Limited computational resources may restrict the scale and complexity of scenarios. This will be addressed through efficient algorithm design and selective focus on most relevant scenarios.

Generalizability: Results from a single intersection may not fully generalize to network-wide applications. This limitation will be explicitly discussed, and future research directions will be identified.

Despite these limitations, the evaluation will provide valuable insights into the potential of AI-powered traffic control in Kenyan urban environments and establish a foundation for further research and development.

Chapter 7

Work Plan and Conclusion

7.1 Detailed Work Plan

The research will be conducted over a 12-month period, following the detailed work plan outlined in Table 7.1. The plan includes specific activities, deliverables, and milestones for each phase of the research.

Timeframe	Activities and Tasks	Deliverables
Month 1	**Project Initiation and Literature Review**	
	• Finalize research proposal and obtain approvals	Research Proposal
	• Conduct comprehensive literature review	Annotated Bibliography
	• Define detailed technical requirements	Technical Specification
	• Set up development environment	Development Environment
Month 2	**Theoretical Framework Development**	
	• Develop integrated theoretical framework	Theoretical Framework Document
	• Define evaluation metrics and methodology	Evaluation Framework
	• Identify simulation scenarios and parameters	Scenario Definition

	• Prepare ethics considerations documentation	Ethics Documentation
Months 3-4	**Simulation Environment Development**	
	• Create digital twin of representative intersection	Functional Simulation Environment
	• Model vehicle types and behavior parameters	Vehicle Behavior Models
	• Develop baseline traffic scenarios	Baseline Scenarios
	• Implement data collection mechanisms	Data Collection System
Months 5-6	**AI Module Development - Phase 1**	
	• Implement synthetic data generation pipeline	Synthetic Dataset
	• Develop and train perception module	Trained Perception Models
	• Implement vehicle detection and tracking	Object Detection System
	• Conduct initial module testing	Module Test Reports
Months 7-8	**AI Module Development - Phase 2**	
	• Implement prediction module with LSTM	Traffic Prediction System

	• Develop reinforcement learning controller	Adaptive Control Algorithm
	• Integrate AI modules with simulation	Integrated Framework
	• Conduct integration testing	Integration Test Reports
Months 9-10	**Comprehensive Evaluation**	
	• Execute evaluation experiments	Experimental Results
	• Collect and analyze performance data	Performance Analysis
	• Conduct statistical analysis	Statistical Report
	• Perform sensitivity analysis	Sensitivity Analysis Report
Month 11	**Results Interpretation and Refinement**	
	• Interpret evaluation results	Results Interpretation
	• Refine models based on findings	Refined Models
	• Prepare preliminary conclusions	Preliminary Conclusions
	• Draft research publications	Manuscript Drafts
Month 12	**Thesis Preparation and Submission**	
	• Write complete thesis document	Final Thesis

	• Prepare final research publications	Submission-ready Manuscripts
	• Develop implementation guidelines	Implementation Guidelines
	• Prepare project documentation	Complete Documentation

Table 7.1: Detailed Work Plan

7.2 Resource Requirements

The research requires several types of resources for successful implementation:

7.2.1 Computational Resources

- **High-Performance Workstation:** Computer with multi-core processor, 32GB+ RAM, and GPU support for deep learning tasks.
- **Software Licenses:** primarily open-source software (SUMO, Python libraries) to minimize costs.
- **Storage:** Adequate storage for simulation data, model files, and results (1TB+ recommended).

7.2.2 Data Resources

- **Reference Data:** Available traffic data from Kenyan authorities for model calibration.
- **Literature Resources:** Access to academic databases and publications.

7.2.3 Human Resources

- **Supervision:** Guidance from academic supervisors.
- **Technical Expertise:** Consultation with domain experts as needed.

7.3 Risk Assessment and Mitigation

Several risks could impact the successful completion of the research. These risks and their mitigation strategies are outlined in Table 7.3.

Risk Category	**Description**	**Mitigation Strategy**	**Probability**
Technical Risks	Simulation environment fails to accurately represent real-world conditions	Careful calibration using available data; sensitivity analysis; expert validation	Medium
	AI models fail to converge or perform poorly	Implement multiple algorithms; extensive hyperparameter tuning; use established architectures	Low
	Integration challenges between modules	Modular design; comprehensive interface specifications; early integration testing	Medium
Resource Risks	Insufficient computational resources for large-scale simulations	Optimize code efficiency; use cloud resources if needed; focus on most critical scenarios	Low
	Data scarcity limits model training and validation	Leverage synthetic data generation; data augmentation techniques; transfer learning	Medium

Timeline Risks	Delays in specific tasks affecting overall timeline	Buffer time in schedule; parallel task execution where possible; regular progress monitoring	Medium
	Scope creep expanding beyond original objectives	Clear scope definition; regular scope reviews; phased implementation approach	Low
Quality Risks	Results not meeting academic standards	Regular supervision meetings; peer review; adherence to research methodology	Low
	Limited generalizability of findings	Explicit discussion of limitations; recommendations for future research	Medium

7.4 Dissemination Plan

The research findings will be disseminated through multiple channels to maximize impact:

7.4.1 Academic Dissemination

- **Master's Thesis:** Comprehensive document submitted to Egerton University.
- **Journal Articles:** Submission to reputable journals in transportation and computer science.
- **Conference Presentations:** Presentation at relevant academic conferences.

7.4.2 Practical Dissemination

- **Technical Reports:** Reports tailored for traffic authorities and policymakers.

- **Open-Source Release:** Public release of the simulation framework and code.
- **Workshops:** Workshops for stakeholders in Kenyan traffic management.

7.4.3 Public Dissemination

- **Summary Briefs:** Non-technical summaries for general audience.
- **Media Engagement:** Engagement with media to share key findings.

7.5 Conclusion

This research proposal presents a comprehensive framework for evaluating AI-powered traffic control strategies in Kenyan urban environments through simulation. The proposed approach addresses the significant challenge of traffic congestion in rapidly growing cities like Nairobi and Nakuru, where conventional traffic management methods have proven inadequate.

The research makes several important contributions. First, it develops a simulation-based methodology that overcomes the data scarcity challenges typical in developing contexts through synthetic data generation. Second, it creates an integrated framework that combines perception, prediction, and control modules using state-of-the-art AI techniques. Third, it provides a rigorous evaluation methodology for comparing AI-powered strategies with conventional approaches across multiple performance dimensions.

The expected outcomes include an open-source simulation framework, validated AI models tailored to Kenyan conditions, performance benchmarks, and practical implementation guidelines. These outputs will contribute to both academic knowledge and practical applications in traffic management.

The research is feasible within the proposed timeline and resource constraints, with a clear work plan and risk mitigation strategies. The findings have the potential to inform traffic management policies in Kenya and similar contexts, demonstrating how simulation-driven approaches can bridge the gap between theoretical research and practical implementation in resource-constrained environments.

By leveraging advanced AI techniques within a simulation framework specifically designed for Kenyan conditions, this research represents an important step toward more intelligent, adaptive, and effective traffic management systems that can address the pressing challenges of urban mobility in developing cities.

Bibliography

Abduljabbar, R., Dia, H., and Tsai, P.-W. (2021). Artificial intelligence in transportation systems: A review. *IEEE Access*, 9:101608–101628.

Bank, A. D. (2021). Urban mobility challenges in african cities. *African Economic Outlook*, 2021:156–178.

Barceló, J. (2010). Fundamentals of traffic simulation. *International Series in Operations Research & Management Science*, 145:1–62.

Cheng, J., Chen, W., Tao, F., and Lin, C.-L. (2020). Digital twin technology for transportation systems: A review. *IEEE Transactions on Intelligent Transportation Systems*, 22(7):4061–4075.

Chu, T., Wang, J., and Chen, L. (2019). Reinforcement learning for traffic signal control: A review and recent advances. *IEEE Transactions on Intelligent Transportation Systems*, 20(12):4406–4424.

Cooper, E. and Barlow, S. (2021). Environmental impact assessment of traffic using simulation. *Transportation Research Part D: Transport and Environment*, 92:102723.

Gao, K., Zhang, Y., Su, R., and Li, L. (2020). Reinforcement learning for traffic signal control: A comparative study. *IEEE Transactions on Intelligent Transportation Systems*, 22(1):286–297.

Ghaffarian, S., Taghavi, M., and Asadi, S. (2021). Digital twins and smart cities: Current status and future trends. *Sustainable Cities and Society*, 75:103301.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.

Guo, Q., Angah, O., and Liu, Z. (2020). Computer vision for traffic monitoring: A comprehensive review. *IEEE Transactions on Intelligent Transportation Systems*, 22(4):1994–2014.

Helbing, D. (2001). Traffic and related self-driven many-particle systems. *Reviews of modern physics*, 73(4):1067.

Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 28(1):75–105.

Kamau, D. and Ochieng, F. (2020). Economic cost of traffic congestion in african cities: A case study of nairobi. *African Development Review*, 32(3):345–361.

Kaplan, A. and Haenlein, M. (2021). Digital twins in smart cities: A systematic review. *Business Horizons*, 64(5):667–679.

Kerner, B. S. (2004). Three-phase traffic theory and highway capacity. *Physica A: Statistical Mechanics and its Applications*, 333:379–440.

Kleijnen, J. P. (2015). Design and analysis of simulation experiments. *International workshop on simulation*, pages 3–22.

Kumar, A. and Barrett, F. (2020). Informal public transport and urban mobility in developing countries. *Transport Reviews*, 40(2):216–234.

Lana, I., Del Ser, J., Velez, M., and Vlahogianni, E. I. (2021). Machine learning for traffic prediction: A systematic review. *Neural Computing and Applications*, 33(15):8947–8975.

Law, A. M. (2015). *Simulation modeling and analysis*. McGraw-Hill Education.

Li, Y., Ouyang, W., Zhou, B., and Wang, X. (2016). Deep learning for traffic perception: A comprehensive review. *IEEE Transactions on Intelligent Transportation Systems*, 18(6):1424–1445.

Liu, Y., Lai, X., and Chang, G.-L. (2020). Emergency evacuation planning using simulation: A review. *Safety science*, 129:104789.

Lopez, P., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., et al. (2018). Microscopic traffic simulation using sumo. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 2575–2582. IEEE.

Maitha, J. and Nyong'o, B. (2020). Institutional challenges in traffic management: The kenyan perspective. *East African Journal of Public Policy*, 8(1):112–128.

May, A. D. (1990). *Traffic Flow Fundamentals*. Prentice Hall.

Mohan, D. and Tiwari, G. (2017). Traffic management in developing cities: Lessons from india. *Transportation Research Part A: Policy and Practice*, 105:64–78.

of Statistics, K. N. B. (2019). Kenya population and housing census. Technical report, Government of Kenya.

Otieno, S., Mwangi, G., and Kamau, D. (2021). Traffic congestion in nairobi: Economic impacts and mitigation strategies. *Journal of Urban Transportation Studies*, 15(2):45–62.

Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., and Wang, Y. (2003). Review of road traffic control strategies. *Proceedings of the IEEE*, 91(12):2043–2067.

Redmon, J. and Farhadi, A. (2018). Yolov3: An incremental improvement. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–6.

Salazar, M., Rodriguez, C., and Lopez, J. (2021). Traffic simulation for developing cities: Challenges and opportunities. *Journal of Advanced Transportation*, 2021:1–15.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Transport, N. and Authority, S. (2022). National transport and safety authority annual report. Technical report, Government of Kenya.

Vapnik, V. N. (1999). *The nature of statistical learning theory*. Springer Science & Business Media.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Vlahogianni, E. I., Karlaftis, M. G., and Golias, J. C. (2014). Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies*, 43:3–19.

Wei, H., Zheng, G., Yao, H., and Li, Z. (2019). Multi-agent reinforcement learning for traffic signal control: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 21(5):1836–1849.

Zhang, J., Wang, F.-Y., and Wang, K. (2018). Intelligent transportation systems: A comprehensive survey. *IEEE Transactions on Intelligent Transportation Systems*, 19(11):3454–3476.

Zhang, L., Liu, Q., Yang, X., and Qiu, T. Z. (2019). Traffic simulation for urban planning: A review. *Journal of Traffic and Transportation Engineering (English Edition)*, 6(2):111–126.

Zhao, Z., Chen, W., Wu, X., Chen, P. C. Y., and Liu, J. (2020). Lstm network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 11(2):68–75.

Appendix A

SUMO Configuration Example

This appendix provides an example SUMO configuration file for the simulation environment.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<configuration xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:
    noNamespaceSchemaLocation="http://sumo.dlr.de/xsd/sumoConfiguration.xsd">
    <input>
        <net-file value="network.net.xml"/>
        <route-files value="routes.rou.xml"/>
        <additional-files value="detectors.add.xml"/>
    </input>
    <time>
        <begin value="0"/>
        <end value="3600"/>
        <step-length value="1"/>
    </time>
    <processing>
        <ignore-route-errors value="true"/>
        <time-to-teleport value="300"/>
    </processing>
    <report>
        <verbose value="true"/>
        <no-step-log value="true"/>
    </report>
</configuration>
```

Listing A.1: Example SUMO Configuration File

Appendix B

Vehicle Type Definitions

This appendix defines the vehicle types used in the simulation, representing the typical vehicle mix in Kenyan urban traffic.

```xml
<!-- Personal Cars -->
<vType id="car" accel="2.6" decel="4.5" sigma="0.5" length="4.5" minGap="2.5" maxSpeed="55" color="1,0,0"/>

<!-- Matatus (14-seater) -->
<vType id="matatu_14" accel="2.4" decel="4.0" sigma="0.6" length="6.0" minGap="2.5" maxSpeed="50" color="0,1,0"/>

<!-- Matatus (25-seater) -->
<vType id="matatu_25" accel="2.2" decel="3.8" sigma="0.6" length="8.0" minGap="2.5" maxSpeed="50" color="0,0,1"/>

<!-- Motorcycles -->
<vType id="motorcycle" accel="3.0" decel="6.0" sigma="0.7" length="2.0" minGap="1.5" maxSpeed="60" color="1,1,0"/>

<!-- Buses -->
<vType id="bus" accel="1.8" decel="3.5" sigma="0.5" length="12.0" minGap="3.0" maxSpeed="45" color="1,0,1"/>

<!-- Trucks -->
<vType id="truck" accel="1.5" decel="3.0" sigma="0.5" length="16.0" minGap="3.0" maxSpeed="40" color="0,1,1"/>
```

Listing B.1: Vehicle Type Definitions

Appendix C

Evaluation Metrics Calculation

This appendix provides the mathematical formulas for calculating the key performance metrics used in the evaluation.

C.1 Efficiency Metrics

C.1.1 Average Delay per Vehicle

$$D_{avg} = \frac{1}{N} \sum_{i=1}^{N} (t_{i,actual} - t_{i,freeflow}) \tag{C.1}$$

Where:

- N: Total number of vehicles
- $t_{i,actual}$: Actual travel time of vehicle i
- $t_{i,freeflow}$: Free-flow travel time of vehicle i

C.1.2 Throughput

$$Q = \frac{N}{T} \times 3600 \tag{C.2}$$

Where:

- N: Number of vehicles passing through intersection
- T: Observation period in seconds

C.2 Environmental Metrics

C.2.1 Fuel Consumption

Fuel consumption is estimated using the following formula based on vehicle kinematics:

$$FC = \sum_{i=1}^{N} \int_{0}^{T} (a \cdot v_i(t) + b \cdot v_i(t)^2 + c \cdot v_i(t)^3 + d \cdot a_i(t)^2)dt \tag{C.3}$$

Where:

- $v_i(t)$: Speed of vehicle i at time t
- $a_i(t)$: Acceleration of vehicle i at time t
- a, b, c, d: Vehicle-specific parameters

C.3 Safety Metrics

C.3.1 Time-to-Collision

$$TTC_{ij} = \frac{x_j - x_i - L_i}{v_i - v_j} \quad \text{for} \quad v_i > v_j \tag{C.4}$$

Where:

- x_i, x_j: Positions of leading and following vehicles
- L_i: Length of leading vehicle
- v_i, v_j: Speeds of following and leading vehicles