Comprehensive Transportation Equity Analysis for RITI Community: A Data-Driven Approach with Case Study

FINAL PROJECT REPORT

by

Yinhai Wang, Wei Sun, Sam Ricord, Cesar Maia de Souza, Yifan Ling, Yan Shi, Bingzhang Wang, Mehrdad Nasri
University of Washington

for

Center for Safety Equity in Transportation (CSET)
USDOT Tier 1 University Transportation Center
University of Alaska Fairbanks
ELIF Suite 240, 1764 Tanana Drive
Fairbanks, AK 99775-5910

In cooperation with U.S. Department of Transportation,
Research and Innovative Technology Administration (RITA)

rural • isolated • tribal • indigenous

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The Center for Safety Equity in Transportation, the U.S. Government and matching sponsor assume no liability for the contents or use thereof.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Comprehensive Transportation Equity Analysis for RITI Community: A Data-Driven Approach with Case Study		**5. Report Date** Jun 30, 2024 **6. Performing Organization Code**
7. Author(s) and Affiliations Yinhai Wang, Wei Sun, Sam Ricord, Cesar Maia de Souza, Yifan Ling, Yan Shi, Bingzhang Wang, Mehrdad Nasri		**8. Performing Organization Report No.** INE/CSET 24.11
9. Performing Organization Name and Address Center for Safety Equity in Transportation ELIF Building Room 240, 1760 Tanana Drive Fairbanks, AK 99775-5910		**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.**
12. Sponsoring Organization Name and Address United States Department of Transportation Research and Innovative Technology Administration 1200 New Jersey Avenue, SE Washington, DC 20590		**13. Type of Report and Period Covered** Final report, Sep 2021 – Jun 2024 **14. Sponsoring Agency Code**
15. Supplementary Notes Report uploaded to:		

16. Abstract

The report presents a comprehensive analysis addressing safety equity concerns for rural communities and American Indian and Alaskan Native (AI/AN) populations, who experience disproportionate rates of serious injuries, fatalities, and general collisions. Despite these disparities, significant gaps exist in understanding the demographics of collisions, particularly within tribal communities where law enforcement jurisdictions are complex, and individuals may misreport their tribal status to gain benefits, leading to biases in collision data. This study aims to fill these gaps by developing a statistical model to predict the true demographics of collisions and enhance safety equity. An ecological regression model, accounting for individual-level characteristics influencing collision rates, is employed. Focusing on Yakima County, Washington—a rural area with the large Yakama Nation reservation—the study examines the impact of household income and AI/AN status on collision rates across three categories: all collisions, injury collisions, and fatal collisions. The results reveal that lower-income individuals are slightly overrepresented in collisions, while higher-income individuals are underrepresented. However, AI/ANs are significantly overrepresented in all collision types, being 3.8 times more likely to be involved in fatal collisions compared to the general population. These findings highlight the utility of ecological regression in revealing the true demographics of collisions and underscore critical safety equity issues in rural and AI/AN communities.

17. Key Words Safety Data Tool; Roadway Safety Assessment; Pedestrian safety, collisions, severity, HSIS database, Machine Learning, Statistical and Machine Learning Modeling, Classification methods;		**18. Distribution Statement**	
19. Security Classification (of this report) Unclassified.	**20. Security Classification (of this page)** Unclassified.	**21. No. of Pages** 35	**22. Price** N/A

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized.**

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Safety equity is a critical concern for rural communities and American Indian and Alaskan Native (AI/AN) populations, who face disproportionate rates of serious injuries, fatalities, and general collisions. Despite these disparities, there are significant gaps in understanding the demographics of collisions, particularly in tribal communities where law enforcement jurisdictions are complex, and individuals may misreport their tribal status to gain benefits, leading to biases in collision data. This study addresses these gaps by developing a statistical model that predicts the true demographics of collisions to enhance safety equity. The ecological regression model is employed, which accounts for individual-level characteristics influencing collision rates. Focusing on Yakima County, Washington, a rural area with a large reservation, the Yakama Nation, this study examines the impact of household income and AI/AN status on collision rates across three collision categories: all collisions, injury collisions, and fatal collisions. Results indicate that lower-income individuals are slightly overrepresented in collisions, while higher-income individuals are underrepresented. AI/ANs, however, are significantly overrepresented in all collision types, being 3.8 times more likely to be involved in fatal collisions compared to the general population. These findings demonstrate the utility of ecological regression in revealing the true demographics of collisions and highlight critical safety equity issues in rural and AI/AN communities.

CHAPTER 1. INTRODUCTION

1.1. Background

Safety is a critical concern in the field of transportation. Over time, traffic fatalities and serious injuries have continued to rise in the United States (Washington Traffic Safety Commission 2019). This trend is especially true in American Indian and Alaskan Native (AI/AN) and rural communities. Both of these areas see a disproportionate rate of fatal and serious injury collisions when compared to both the national average and nearby urban areas. Fatality rates are nearly double for those living in rural communities (Federal, 2012). AI/ANs are also nearly four and a half times as likely to be killed in a fatal collision than non-native populations in the state of Washington (Washington,2019). These disparities highlight a significant equity concern due to the disproportionate distribution of collisions to these historically disadvantaged communities.

The root causes of these disparities are multifaceted and complex, involving a combination of socio-economic, infrastructural, and behavioral factors. Rural areas often suffer from underfunded and poorly maintained road infrastructure, which can lead to hazardous driving conditions. Additionally, emergency response times in these areas are generally longer due to the greater distances that need to be covered. This delay can be critical in determining the survival and recovery outcomes of those involved in serious collisions. The socio-economic context also plays a significant role, as individuals in rural and tribal communities may have less access to newer, safer vehicles and may lack the resources for regular vehicle maintenance.

Furthermore, the behavioral aspects of traffic safety in rural and tribal communities cannot be ignored. Higher incidences of risky driving behaviors, such as speeding, driving under the influence of alcohol or drugs, and lower seatbelt usage rates, are often reported in these areas. These behaviors are exacerbated by the limited presence of law enforcement and traffic safety campaigns that are more prevalent in urban areas. The cultural context within AI/AN communities also adds layers of complexity, where traditional practices and beliefs may influence driving behaviors and attitudes towards traffic safety regulations.

Despite these observed disparities in safety equity for rural and tribal communities, there are still significant gaps in the understanding of the demographics of collision victims for these communities. Tribal communities especially encounter significant jurisdictional issues related to the recording of demographics in collisions. Tribal communities on and around reservations already have significantly more jurisdictional crossover and confusion in law enforcement, having to account for not only city, county, and state patrol law enforcement but also tribal police, Bureau of Indian Affairs (BIA) law enforcement, and the Federal Bureau of Investigation (FBI). This already leads to more confusion and errors in the reporting of collisions as each law enforcement agency has different policies and historical behavior related to duties of traffic enforcement and collision response. Beyond this, the implications of dealing with each of these different law enforcement groups for a tribal member can vary widely depending on the context. Therefore, it is not uncommon for tribal members to selectively report their membership to the tribe only when the context of the particular incident would be in their favor. This leads to further significant biases in the demographic data related to collisions in tribal communities where there is no guarantee that tribal members will accurately report their true demographics.

The jurisdictional challenges are compounded by historical mistrust between tribal communities and external law enforcement agencies. This mistrust stems from a long history of discrimination, marginalization, and broken treaties, leading to a lack of cooperation in data sharing and enforcement of traffic laws. Moreover, the fragmentation of law enforcement responsibilities often results in inconsistent data collection methods and standards, further affecting the accuracy of collision data (Smith, 2023).

This paper addresses this key gap in understanding uncertain collision demographics by utilizing statistical methods to predict the true demographics of those involved in collisions in a rural and tribal community. The statistical methodology used in this paper is called ecological regression. This technique can utilize the individual level characteristics of collision data to extract a true estimation of the demographics described by the data (Jackson, 2006). This paper conducts a case study in Yakima County of Washington State. This county has a major tribal community present, as it is home to the majority of the Yakama Nation Reservation. This study will help address the key safety equity issue by providing better understanding of the demographics involved in collisions which can help to better target the most pressing safety issues faced by these communities through the four E's of traffic safety: Engineering, Education, Enforcement, and Emergency Services.

Engineering interventions can include the redesign of hazardous road segments, the installation of better signage and lighting, and the implementation of traffic calming measures to reduce speeds. Education efforts could focus on culturally sensitive outreach programs that resonate with AI/AN communities, promoting safe driving practices and increasing awareness of traffic laws. Enforcement strategies might involve enhancing the presence and capabilities of tribal law enforcement agencies and fostering cooperation with other jurisdictions to ensure consistent application of traffic safety regulations. Finally, improving Emergency Services could mean better training for first responders on handling collisions in tribal areas and ensuring faster response times to these often-remote locations.

The case study in Yakima County aims to serve as a model for other regions with similar demographic and jurisdictional challenges. By providing a clearer picture of the true demographics involved in traffic collisions, policymakers and community leaders can develop targeted interventions that address the unique needs of rural and AI/AN populations. This approach not only aims to reduce traffic fatalities and injuries but also strives to rectify long-standing inequities in traffic safety.

In conclusion, addressing the safety disparities in rural and tribal communities requires a comprehensive understanding of the underlying factors contributing to higher fatality and injury rates. Through the application of advanced statistical methods and a focus on the four E's of traffic safety, this paper seeks to provide actionable insights that can inform better policy decisions and ultimately save lives. The findings from Yakima County can pave the way for broader initiatives aimed at improving traffic safety equity for all underserved communities across the United States.

1.2. Practical Application

Safety equity remains a critical issue for rural and AI/AN communities, both of which face disproportionately high rates of collisions. This study contributes to addressing this issue by presenting a method to estimate the true demographics of individuals involved in traffic collisions through statistical modeling. This approach allows safety practitioners to better understand which demographic groups are more likely to be involved in collisions, thereby providing valuable insights to address safety equity issues faced by disadvantaged communities. The methodology can be applied to a wide range of demographic categories, including income, race, and more.

The study specifically examines the AI/AN population in Yakima County, Washington, revealing that these populations are approximately 3.8 times more likely to be killed in a fatal collision compared to the general population. This finding underscores the importance of targeted safety interventions. The ability to accurately predict collision demographics can guide the allocation of resources and the development of policies aimed at reducing collision rates among overrepresented groups. For example, implementing targeted educational campaigns, enhancing law enforcement efforts, and improving infrastructure in areas with high rates of collisions can significantly impact safety outcomes.

In practice, the methodology outlined in this study can be expanded to other communities across the state of Washington and beyond. By applying ecological regression to various regions, researchers and policymakers can gain a more comprehensive understanding of the safety and equity concerns faced by different communities. This broader application can identify critical safety issues and guide the implementation of targeted interventions to achieve safety equity. Moreover, the study highlights the necessity of addressing jurisdictional complexities in tribal areas. Improved coordination between different law enforcement agencies and standardized reporting practices can reduce biases in collision data, leading to more accurate demographic profiles and better-targeted safety measures. This approach aligns with the four E's of traffic safety: Engineering, Education, Enforcement, and Emergency Services, providing a holistic framework to address safety equity issues.

The results of this study also have significant implications for Vision Zero initiatives, which aim to eliminate all serious injuries and fatalities on roadways. By understanding the true demographics of collisions, Vision Zero strategies can be more effectively tailored to address the specific needs of high-risk groups, particularly in rural and tribal communities. This targeted approach can help bridge the gap in safety equity and contribute to the overall goal of reducing traffic-related fatalities and injuries. The methodology can also be adapted to other historically disadvantaged communities, providing a valuable tool for researchers and policymakers to address safety disparities. By utilizing advanced statistical methods, such as machine learning and artificial intelligence, the predictive accuracy and explanatory power of collision models can be further enhanced, leading to more effective safety interventions.

In conclusion, this study provides a robust framework for understanding and addressing safety equity in rural and AI/AN communities. By accurately predicting the demographics of collisions, policymakers can implement targeted interventions to reduce disparities and improve overall traffic safety. The application of ecological regression and other advanced statistical methods represents a significant advancement in achieving safety equity and ensuring that all communities benefit from safe and equitable transportation systems.

CHAPTER 2. LITERATURE REVIEW

To address the issue of unknown demographics in collision reporting and to understand the impact that it has on equity in traffic safety, it is crucial to first define equity explicitly within this context. Transportation, as a multifaceted domain, evokes a myriad of definitions and interpretations of equity, each with its nuances and implications. The concept of equity in transportation can be approached from various theoretical perspectives, each proposing different criteria for what constitutes a 'just' distribution of resources and outcomes. One such comprehensive framework is provided by Lewis et al. (2021), who discuss the plurality of equity theories. According to Lewis, there are fundamentally different, and sometimes conflicting, theories on how resources should be distributed to achieve justice. This divergence in perspectives necessitates a clear and context-specific definition of equity, particularly in traffic safety, where the stakes involve both human lives and societal well-being (Lewis et al., 2021).

A notable study by Martens and Golub (2021) mirrors this idea by proposing a normative framework for addressing equity in transportation. They introduce a four-rung ladder that utilizes different interpretations of Title VI of the Civil Rights Act to define equity. The rungs of the ladder, in ascending order, are: Explicit Non-Discrimination, Pareto-Plus Improvement, Proportional Equity, and Restorative Justice (Martens and Golub 2021). The first rung, Explicit Non-Discrimination, mandates that transportation policies and practices should not explicitly discriminate against any demographic group. This is the most basic form of equity, ensuring that no overt bias exists in the allocation of resources. The second rung, Pareto-Plus Improvement, goes a step further by suggesting that changes in the transportation system should benefit at least one person without making anyone else worse off. This principle, while still conservative, begins to acknowledge the need for improvements that consider the well-being of all. Proportional Equity, the third rung, argues for a distribution of resources such that the intended outcomes are equally shared among the entire population. This rung emphasizes equal access and benefits, ensuring that transportation services do not disproportionately favor or disadvantage any particular group. The fourth and highest rung, Restorative Justice, incorporates the principles of Proportional Equity but extends them to address historical injustices. This involves not only equal distribution but also compensating for historically denied benefits, thereby attempting to rectify past inequities and provide a more comprehensive form of justice. These differing interpretations underscore the complexity of defining and achieving equity in transportation. The ladder framework elucidates the transition from basic non-discrimination to a more nuanced and historically aware understanding of justice (Martens and Golub 2021).

The implications of these concepts are profound, as transportation plays a vital role in societal functioning by providing access to essential goods and services. Pereira et al. (2021) highlights that the historical and ongoing inaccessibility of these resources can significantly hinder community growth and development. Addressing these inequities through the lens of justice aims to rectify past wrongs and promote a more equitable society. Hence, defining equity in transportation requires a multifaceted approach that considers both present and historical contexts (Pereira et al., 2021). By adopting a comprehensive framework, policymakers can better address the nuances of equity and work towards a more just distribution of transportation resources and outcomes.

In many instances, local implementations of equity in transportation have produced varying results, reflecting the diverse methods used to define and quantify equity (Lewis et al., 2021). These variations highlight the complexity and challenges of achieving equitable outcomes. While many current municipalities employ some form of aggregate equity metrics to guide their projects, these metrics often prove inadequate. Aggregate metrics can obscure significant variations within demographic groups and fail to account for the differing needs of various populations in terms of access to resources (Martens and Golub, 2022). This shortfall can lead to suboptimal and inequitable distributions of resources.

A prime example of this issue is illustrated in a study conducted in California. The study found that Metropolitan Planning Organizations (MPOs), when attempting to enhance public transit services, typically focused on increasing transit availability within urban cores (Heyer et al, 2020). This strategy was chosen because it promised the greatest aggregate benefits, such as increased ridership, reduced vehicle miles traveled, and lowered air pollution. However, this approach systematically neglected some of the poorest communities. These communities, driven to the fringes of metropolitan areas by rising housing costs, were underserved by transit improvements concentrated in urban centers. Consequently, the equity goals of these initiatives were undermined, exacerbating existing inequalities rather than alleviating them (Heyer et al., 2020).

Addressing these challenges requires more sophisticated methodologies that can capture the nuances of equity. One such approach is the robust decision-making (RDM) method. RDM is designed to handle deep uncertainty and is particularly useful in planning for equitable outcomes. It involves evaluating a wide range of scenarios to identify strategies that are robust across different possible futures. By considering numerous potential outcomes, RDM aims to ensure that equity goals are met more consistently and effectively, even under varying conditions (Lempert et al., 2020).

These issues are not confined to transit equity but are also prevalent in traffic safety. The interpretation of equity in traffic safety can vary, encompassing multiple aspects of the equity ladder described earlier. For the purposes of this study, the focus will be on the top two rungs: proportional equity and restorative justice, which are widely recognized as standards for transportation equity (Martens and Golub 2021). In traffic safety, proportional equity is achieved by reducing or eliminating disparities in collision rates among different demographic groups. This means that collisions should be proportionately distributed, ensuring no demographic group bears a disproportionate burden of traffic incidents. Achieving proportional equity involves targeted interventions that address specific needs and vulnerabilities of different groups. Restorative justice in traffic safety takes a more ambitious approach. It seeks not only to equalize collision rates but also to address historical injustices that have contributed to current disparities. Restorative justice aims for the complete elimination of serious injuries and fatalities, acknowledging that proportional rates alone do not rectify past inequities. This principle is embodied in the Vision Zero movement, which aspires to eliminate all serious injuries and fatalities on roadways. Vision Zero's goals are intrinsically linked to restorative justice, emphasizing comprehensive safety measures that account for historical and contextual factors (Kim et al., 2020).

This study aims to contribute to this framework by improving the understanding of the demographics involved in collisions. By identifying the most affected groups, more effective and targeted campaigns can be developed to reduce collision rates and severity. This targeted approach is essential for achieving both proportional equity and restorative justice in traffic safety, ensuring that interventions are not only equitable but also just.

The concept of Vision Zero is particularly relevant in this context. Vision Zero advocates for a holistic approach to traffic safety, integrating infrastructure improvements, policy changes, and community engagement to achieve its ambitious goals. This movement recognizes that every traffic death is preventable and that achieving zero fatalities is a moral imperative. By focusing on systemic changes and community-specific strategies, Vision Zero aims to create safer road environments for all users. Implementing Vision Zero principles involves a range of strategies, from redesigning dangerous intersections and improving pedestrian crossings to enforcing traffic laws and promoting safe driving behaviors. These measures are designed to protect the most vulnerable road users, including pedestrians, cyclists, and motorcyclists, who are often disproportionately affected by traffic collisions.

Furthermore, achieving restorative justice in traffic safety necessitates a shift in how we approach and prioritize safety interventions. It requires acknowledging and addressing the historical and systemic factors that have contributed to inequitable safety outcomes. This includes investing in underserved communities, ensuring equitable access to safe transportation options, and actively involving community members in decision-making processes.

In conclusion, addressing equity in transportation and traffic safety requires a multifaceted approach that combines proportional equity and restorative justice. By employing targeted interventions and robust methodologies, policymakers can work towards a more equitable and just transportation system. This study contributes to this effort by enhancing our understanding of collision demographics, enabling more effective and equitable safety campaigns. Through these efforts, we can move closer to the vision of eliminating all serious injuries and fatalities on our roadways, ensuring that safety benefits are distributed fairly across all demographic groups.

CHAPTER 3. METHODOLOGY AND RESULTS FOR THE MAIN PROJECT

3.1. METHODOLOGY

This paper employs ecological regression to statistically analyze the demographics of traffic collisions in Yakima County, Washington State. Situated in the central part of the state, Yakima County lies east of the Cascade Mountains, far removed from the bustling urban centers of the greater Seattle area. The county's primary urban hub is the city of Yakima, which boasts a population just over 90,000 residents. Apart from Yakima, the county is predominantly rural, characterized by several small towns, extensive wilderness areas, and vast tracts of farmland that contribute significantly to the local economy.

Yakima County is also notable for hosting one of the largest American Indian/Alaska Native reservations in Washington State, the Confederated Tribes and Bands of the Yakama Nation (Wikipedia, 2024). The Yakama Nation, a federally recognized tribe under the Camp Stevens Treaty of 1855, currently occupies a reservation that spans an impressive 1.37 million acres (StoryMaps, 2021). The towns of Toppenish and Wapato serve as the primary population centers within the reservation. This reservation is a major agricultural powerhouse, known for producing nearly $2 billion in agricultural products annually, which underscores its vital role in both the local and state economies (StoryMaps, 2021).

Despite its agricultural prosperity, Yakima County faces significant traffic safety challenges. Recent data from the Washington Traffic Safety Commission (2019) indicate that Yakima County has experienced the highest number of fatal roadway collisions and pedestrian fatalities in the state. This alarming trend highlights the urgent need for targeted traffic safety interventions and policy measures to mitigate the risks faced by both residents and visitors.

In addition to these concerns, the unique demographic composition of Yakima County, which includes a substantial population of American Indian/Alaska Native individuals, necessitates a nuanced understanding of traffic safety issues. The diverse population dynamics and varying levels of infrastructure development across urban and rural areas contribute to the complexity of addressing traffic safety in the region. The findings of this paper, which leverage ecological regression techniques, aim to provide insights into the specific demographic factors associated with traffic collisions, thereby informing more effective and equitable traffic safety strategies.

To aid in visualizing the geographic context, a map of Yakima County and the Yakama Nation is included in Figure 3-1. By integrating statistical analysis with geographic visualization, this paper seeks to offer a comprehensive overview of the traffic safety landscape in Yakima County, ultimately contributing to the development of data-driven solutions to enhance roadway safety for all community members.

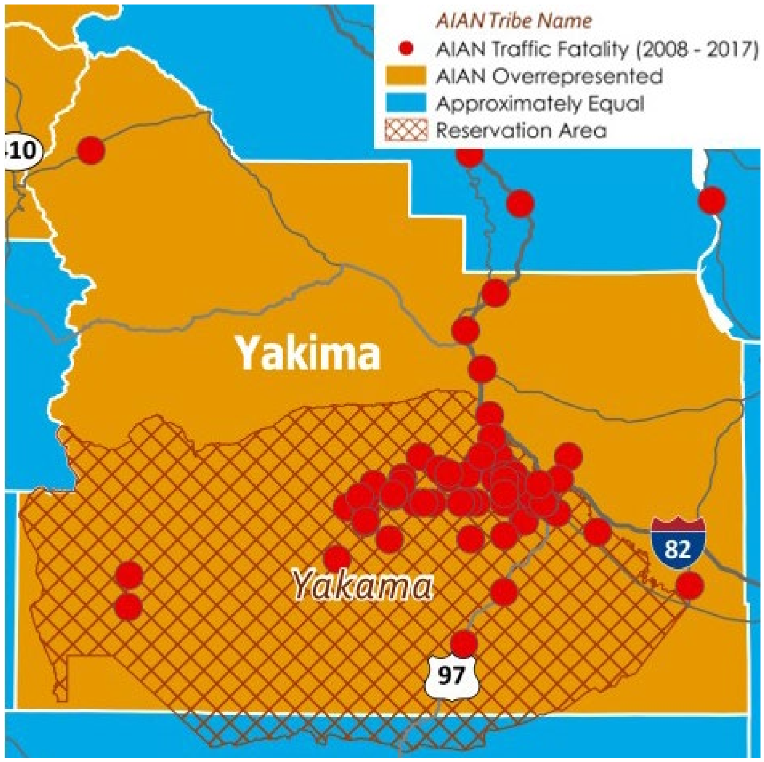


Figure 3-1 Yakima Nation geographic context

Two data sources were utilized in this study. The first source is collision data collected from 2018 to 2021 from the Highway Safety Information System (HSIS), and the second source is Census data collected from the 2020 American Community Survey 1-year estimate. These datasets provide the essential information needed to conduct this study comprehensively. The collision data includes records for most reported collisions across the entire state of Washington, although this study specifically focuses on Yakima County. These collision records are primarily generated through police reports, which can be subject to biases such as those previously mentioned regarding demographic reporting. The second data source is Census data, which offers crucial demographic information, including household income and racial characteristics. Both of these data sources are indispensable for statistically analyzing the demographics involved in collisions. By integrating collision data with detailed demographic information from the Census, the study aims to identify patterns and correlations that may exist between collision occurrences and various demographic factors. This dual-source approach allows for a more nuanced understanding of the factors contributing to collisions in Yakima County, facilitating targeted interventions and policy recommendations.

To facilitate this goal, ecological regression models are employed to predict the true demographics of collisions throughout Yakima County. Ecological regression is an ideal statistical technique for achieving this objective as it can account for individual characteristics to calculate the probabilities that people of different demographic characteristics are represented in transportation datasets (Jackson 2006). This technique can be applied to estimate the demographic profile for various transportation networks by calculating the odds ratios for each demographic strata of interest, determining the probability that each demographic group will be over or undercounted (see Ricord 2023 for more details). In this context, ecological regression will be used to calculate the demographic profiles of the collision records accurately.

This analytical process will be conducted twice: once for household income demographics and once for a binary selector indicating whether a collision involved an AI/AN (American Indian/Alaska Native) individual or not. For each of these demographic categories, three distinct models will be created to analyze the demographics of all collisions, injury collisions, and fatal collisions. This approach necessitates the development of six total models to comprehensively understand the demographics of collisions within the county.

By creating these models, the study aims to provide a detailed analysis of the demographic factors associated with different types of collisions, enabling a deeper insight into the relationship between demographic characteristics and collision occurrences. This multifaceted modeling approach will help identify specific patterns and trends, which can then inform targeted interventions and policy decisions aimed at improving road safety and addressing any identified disparities.

The output of these models will enable the direct estimation of the true demographic profiles of each scenario, specifically the demographics of those involved in the collisions. These models achieve this by linking the locations of crashes to the geographically defined demographics provided by Census data. This approach yields the predicted percentage of crashes associated with each demographic stratum. From these predictions, the over- or under-representation of each stratum is calculated to reveal the true distribution of collisions across different demographics. Consequently, conclusions can be drawn about the actual rate of collisions for various critical demographic groups.

This basic technique has been employed in previous studies to determine the true demographics of collision data. Several studies have utilized Bayesian ecological regression models to analyze road mortality. For instance, a study conducted in Europe aimed to understand regional differences in fatality rates, seeking to derive explanatory results for collisions (Eksler, 2008). Another similar study was conducted in Tunisia (Kammoun 2020). However, this study differs from those in several significant ways. Firstly, this study uses a different type of ecological regression model known as Goodman's Method of Bounds regression. The choice of this model is driven by the differing goals of the studies; while the aforementioned studies aimed to explain collisions based on population density, this study aims to predict collisions based on demographics. Although the fundamental concept of the statistical model is similar, Goodman's Method of Bounds ecological regression models are better suited to this study's specific objectives than Bayesian ecological regression models, and thus were chosen for use in this analysis.

By leveraging Goodman's Method of Bounds, this study aims to provide more precise predictions regarding the demographics involved in collisions. This methodological choice enhances the ability to accurately identify demographic patterns and disparities in collision occurrences. The refined approach ensures that the study's findings are robust and directly applicable to the development of targeted safety interventions and policy recommendations.

3.2. RESULTS

As mentioned above, six ecological regression models were built to determine various demographics of collisions in Yakima County. Table 1 displays the fidelity of all models by presenting the variance of the 95% confidence interval associated with each model. This variance measure provides insight into the precision and reliability of the models.

From Table 3-1, it can be observed that the models predicting the binary involvement rate of AI/AN (American Indian/Alaska Native) individuals in collisions have a smaller confidence interval range compared to those predicting household income demographics. Additionally, the confidence interval variance tends to increase as the type of collision becomes more specific, progressing from all collisions to injury collisions to fatal collisions. Despite this increase, all confidence intervals remain within an acceptable range, indicating the robustness and validity of the models.

Table 3-1 Variance of Each Model's 95% Confidence Interval

	Collision Type		
	All Collisions	Injury Collisions	Fatal Collisions
Household Income	0.00294	0.00538	0.0334
AI/AN	0.000342	0.000585	0.00172

Table 3-2 below presents the odds ratios for the three household income models, while Table 3 shows the odds ratios for the three AI/AN binary models. These odds ratios are crucial for understanding the likelihood of different demographic groups being involved in collisions based on household income and AI/AN status.

Table 3-2 Odds Ratios for Household Income Models

Household Income Bracket	Collision Type All Collisions Odds Ratio	Injury Collisions Odds Ratio	Fatal Collisions Odds Ratio
Intercept	0.160	0.0320	0.0000252
<$10,000	0.999	1.004	1.012
$10,000 to $14,999	1.005	1.004	1.011
$15,000 to $19,999	1.001	0.999	0.971
$20,000 to $24,999	1.001	1.001	1.031
$25,000 to $29,999	1.000	0.999	1.014
$30,000 to $34,999	1.001	1.001	1.010
$35,000 to $39,999	1.002	1.003	0.986
$40,000 to $44,999	0.999	0.998	0.982
$45,000 to $49,999	0.998	0.999	1.011
$50,000 to $59,999	0.999	1.000	1.010
$60,000 to $74,999	0.998	0.999	1.010
$75,000 to $99,999	0.997	0.995	0.996
$100,000 to $124,999	1.003	1.004	1.009
$125,000 to $149,999	1.002	1.003	1.006
$150,000 to $199,999	0.998	0.999	0.995
>$200,000	0.997	0.997	1.008

The tables reveal that the odds ratios for all models are near one. This suggests that the primary factor influencing the demographics of collisions is the geographic distribution of collisions. However, these odds ratios still need to be applied to the collision data to accurately determine the true rate of collision for different demographics. The implications regarding household income are explored first.

Figure 3-2 illustrates the income distributions for each collision type, alongside the income distribution for the entire county. This visual representation helps in comparing the income demographics involved in different types of collisions with the overall county income distribution. Table 4 complements this by showing both the percentage representation of each income stratum in collisions and the likelihood of each stratum being involved in different types of collisions.

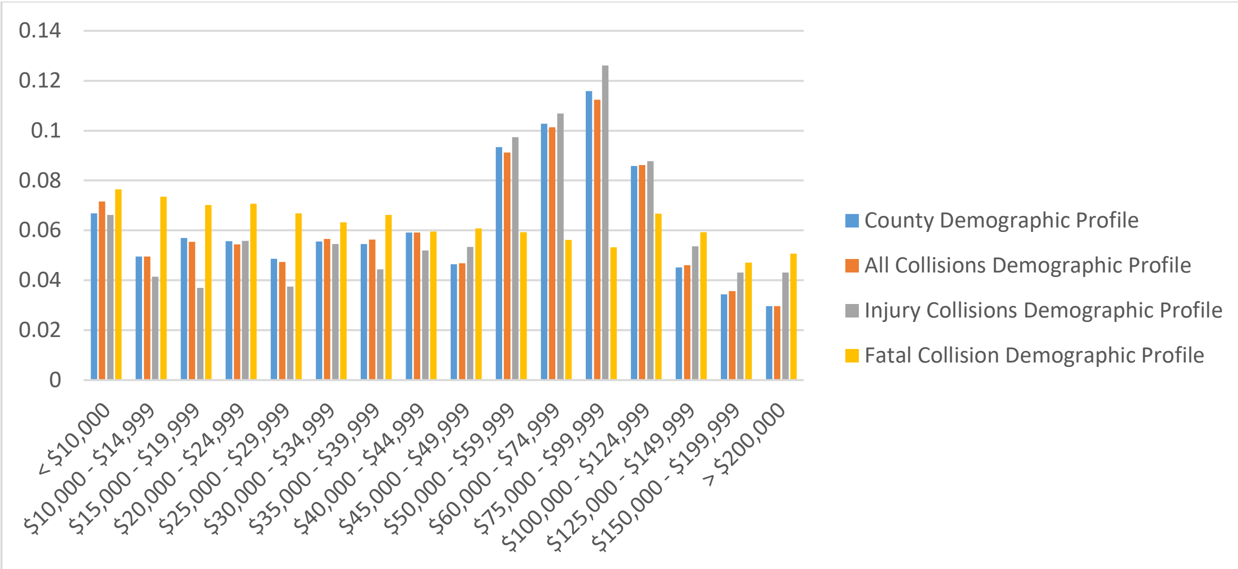


Figure 3-2 Income distribution for each collision type

From these results, several interesting trends can be extracted. Firstly, the demographic profile for all collisions and injury collisions tends to closely follow the general demographic profile of the population. However, there is a noticeable trend where lower-income households are slightly overrepresented in these types of collisions, while higher-income households are slightly underrepresented. This indicates that lower-income households are more frequently involved in all and injury collisions compared to their proportion in the population.

On the other hand, fatal injuries do not follow these trends. Instead, fatal injuries appear to be distributed more evenly among different household income levels, rather than correlating with population size. However, although the percentage of fatal collisions is roughly the same for each demographic stratum, there is a slight decrease in the percentage as household income increases. This suggests that while household income does not significantly correlate with the likelihood of fatal collisions, it does correlate with the likelihood of all and injury collisions. In summary, all types of collisions tend to slightly overrepresent lower-income households and slightly underrepresent higher-income households.

Next, the percentage representation and collision likelihood for AI/AN individuals is examined. In Yakima County, AI/ANs make up 3.52% of the population. However, they account for 4.59% of all crashes, 5.04% of all injury collisions, and a striking 13.37% of all fatal collisions.

These findings indicate that AI/AN individuals are more likely to be involved in collisions compared to the overall population. For all categories, the likelihood of being in a crash is above 1, and this likelihood increases with the severity of the collision. The starkest disparity is observed in fatal collisions, where AI/ANs are nearly four times as likely to be killed in a fatal collision compared to the general population. This significant overrepresentation in fatal collisions highlights a critical area of concern that requires targeted interventions to improve road safety for AI/AN communities.

3.3. CONCLUSIONS

From the results of this study, several conclusions can be drawn. Firstly, this study showcases how ecological regression can be effectively utilized to estimate the true demographic profile of collision data. By linking crash locations with Census data, this method provides a detailed and accurate depiction of the demographics involved in collisions, thereby offering valuable insights into safety equity within Yakima County.

One key finding is that the demographic rates of all collisions and injury collisions correlate closely with the overall demographics of the county. However, there is a noticeable trend of slight overrepresentation of lower-income individuals in these collision types compared to higher-income individuals. This suggests that lower-income households are more frequently involved in both all and injury collisions than their proportion in the population would suggest.

Interestingly, this trend does not extend to fatal collisions. Instead, fatal collisions do not correlate with household income demographics in the same way. The distribution of fatal collisions is more evenly spread across different income brackets. Despite this even distribution, there remains a slight overrepresentation of lower-income households in fatal collisions. This highlights a critical safety equity concern, as lower-income individuals are disproportionately bearing the costs associated with traffic safety.

The overrepresentation of lower-income households in all and injury collisions underscores the need for targeted interventions to improve road safety for these populations. This finding suggests that socioeconomic factors play a significant role in collision involvement, possibly due to differences in vehicle safety, driving conditions, or access to safe infrastructure.

Additionally, the study's results emphasize the importance of addressing safety equity in traffic safety policies and interventions. The fact that lower-income individuals are slightly more impacted by traffic collisions indicates a disparity that needs to be addressed through targeted measures. This could include improving infrastructure in lower-income areas, implementing more stringent safety regulations, and providing better access to safe transportation options for all income groups.

These results have several implications regarding Vision Zero and safety equity. Firstly, they highlight the severe safety equity issues faced by both rural and tribal communities. In this particular rural and tribal community, American Indian/Alaska Native (AI/AN) individuals are nearly 3.8 times as likely to be killed in a fatal collision. This alarming statistic aligns with findings from other studies conducted in the region. According to the Washington Traffic Safety Commission (2019), across the state of Washington, AI/AN individuals are 4.4 times as likely to be killed in a traffic collision. This consistent overrepresentation indicates a pressing need to focus on traffic safety within AI/AN communities, which are disproportionately affected by these safety equity issues.

The disparities are even more pronounced for specific types of collisions. AI/AN individuals are 6.4 times as likely to be killed in a pedestrian collision, 5.8 times as likely to be killed when impairment is involved, 4.2 times as likely to be killed in speeding-related collisions, and a staggering 8.8 times as likely to be killed when the collision involves unrestrained occupants (Washington Traffic Safety Commission 2019). These figures underscore the critical need for targeted interventions and improvements in traffic safety measures for AI/AN communities. To address these disparities and improve safety outcomes, various solutions can be framed around the 4 E's: Engineering, Education, Enforcement, and Emergency Services.

There are several potential avenues by which this work can be built upon to further the goals of safety equity. First and foremost, this research can be expanded to include other communities, both throughout the state of Washington and across the country. By broadening the scope, we can gain a more comprehensive understanding of the demographics involved in collisions, which will enhance our knowledge of the safety and equity concerns faced by diverse communities. This expansion will help to identify the most critical aspects of safety that must be addressed to achieve Vision Zero and realize restorative justice in safety.

Furthermore, extending this study to encompass more tribal communities will provide invaluable insights into the specific safety equity issues these groups face. Given the unique challenges and disparities experienced by tribal communities, targeted research is essential to develop effective interventions. Beyond tribal communities, this methodology can also be applied to other historically disadvantaged communities. By doing so, we can uncover the safety equity issues prevalent in various demographics and work towards addressing these systemic disparities.

In addition to expanding the geographic and demographic scope of the study, there are several ways to improve the methodology itself. One significant improvement could involve the integration of more advanced statistical methods. While ecological regression has proven effective, methods utilizing machine learning and artificial intelligence show great promise for providing deeper explanatory insights into collision and traffic safety data. These advanced techniques can help identify complex patterns and relationships that might not be apparent through traditional statistical methods.

CHAPTER 4. AUXILLARY ROAD SURFACE CONDITION DETECTION TOOL FOR RITI COMMUNITY

Road surface condition detection is a crucial element of transportation safety and infrastructure management. This process employs a variety of technologies and methodologies to identify and evaluate the condition of road surfaces, enabling transportation authorities to maintain safe and efficient roadways. The state of road surfaces directly affects vehicle safety, ride comfort, and overall transportation efficiency. Poor road conditions, such as potholes, cracks, and surface wear, can result in vehicle damage, accidents, and increased travel times. Additionally, adverse weather conditions like ice, snow, and water accumulation can further worsen these risks. Therefore, timely and accurate detection of road surface conditions is essential for preventing accidents by identifying hazardous conditions early, maintaining vehicle integrity by preventing wear and tear, enhancing ride comfort by ensuring smooth surfaces, optimizing travel efficiency by reducing travel time, and managing infrastructure proactively to extend the lifespan of roadways and reduce long-term costs.

- Reducing the risk of collisions caused by poor road conditions to ensure road safety.
- Extending the lifespan of road infrastructure by prioritizing maintenance and repair activities.
- Preventing road closures and delays to enhance traffic flow and reduce congestion.
- Enabling safer driving decisions by informing drivers about current road conditions.

Current methodologies and technologies for detecting road surface conditions include:

- Visual Inspection: This traditional approach involves trained personnel manually inspecting roads. Although straightforward, it is time-consuming, subjective, and limited in scope.
- Sensor-Based Systems: Sensors such as accelerometers, laser scanners, and ultrasonic devices can be mounted on vehicles to detect road surface irregularities during travel. These systems offer continuous and objective data but can be expensive and require complex data analysis.
- Image and Video Analysis: Advances in computer vision and machine learning have enabled the use of cameras to capture and analyze road surface images and videos. Algorithms can detect and classify defects, measure their dimensions, and assess their severity.
- Thermal Imaging: Infrared cameras detect temperature variations on the road surface, indicating the presence of moisture, ice, or snow. This technology is particularly effective for monitoring road conditions in cold climates.

4.1. Methodology

The implementation of AI (Artificial Intelligence) and ML (Machine Learning) algorithms is expected to significantly enhance the accuracy and efficiency of detecting road surface conditions. These technologies enable the processing of large datasets and the recognition of complex patterns. Additionally, the integration of IoT (Internet of Things) devices with road infrastructure can facilitate real-time monitoring and data collection, thereby improving the responsiveness of maintenance efforts.

In our approach to classifying road surface conditions, we have identified four key features: two from image data and two from environmental data. The image-based features are intensity value and dark channel value. According to the Dark Channel Prior (DCP) theory, the dark channel value is usually low for most natural objects, including road surfaces, unless the object is white. However, in rainy or snowy conditions, where reflections or white pixels are common, the dark channel value may be higher. On the other hand, the intensity value is effective in differentiating between snowy and non-snowy conditions, as intensity values are higher in snowy environments. Figures 4-1 and 4-2, showing the intensity and dark channel histograms of road regions in Washington State captured by a surveillance camera, highlight the significant differences between roads with and without snow, which can be leveraged for classification. The addition of temperature and humidity sensors further improves the accuracy and reliability of this classification approach (Liu et al., 2023).

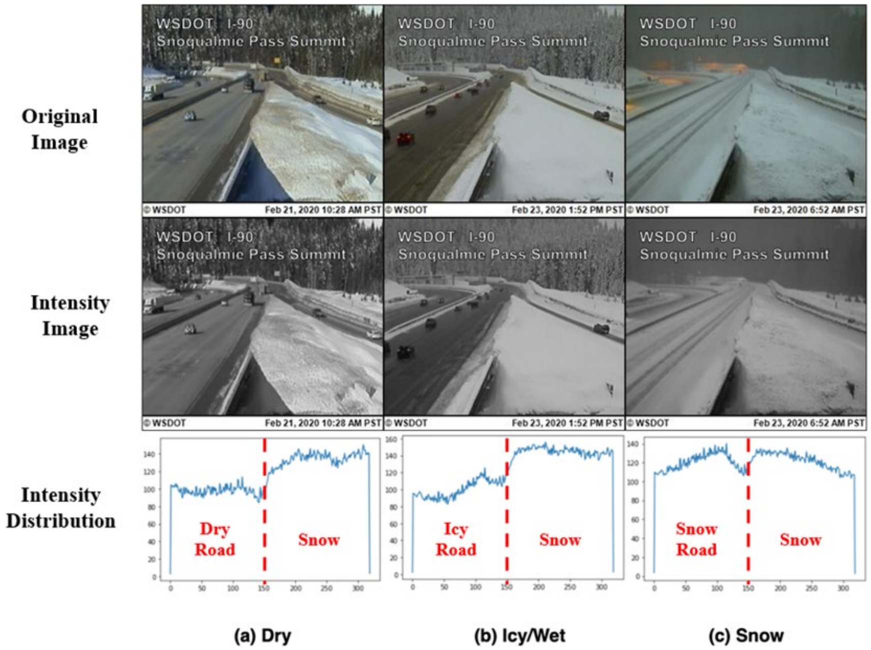


Figure 4-1 Image Intensity of Different Road Surface Conditions (Liu et al. 2023)

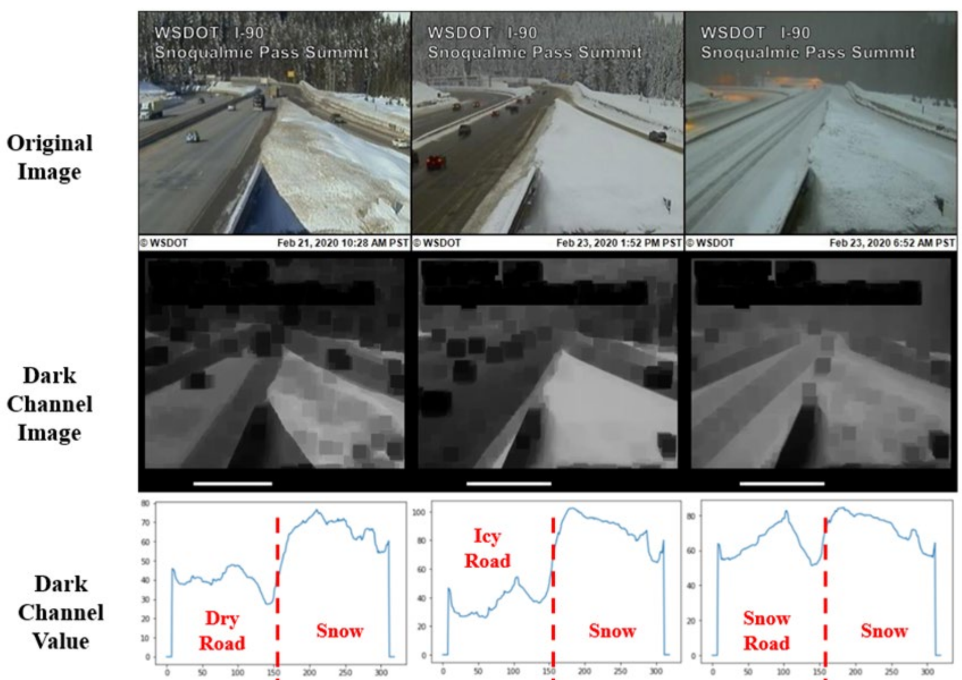


Figure 4-2 Image Dark Channel of Different Road Surface Conditions (Liu et al. 2023)

In addition to image-based features, we also incorporate two environmental features: temperature and humidity, measured by the DHT22 sensor. Temperature acts as an indicator of snowy conditions, while humidity is a signal for rainy conditions. These features are straightforward, representative, and create an effective feature space for classification. The feature vector is represented as [I, K, T, H], where I stands for the median intensity, K denotes the median dark channel value, T represents temperature, and H indicates humidity. Using these four features, COCO SENSOR can classify road surface conditions into categories such as Icy, Snow, Dry, and Wet (Liu et al. 2023). Figure 4-3 illustrates the image processing results of the road surface condition detection tool.

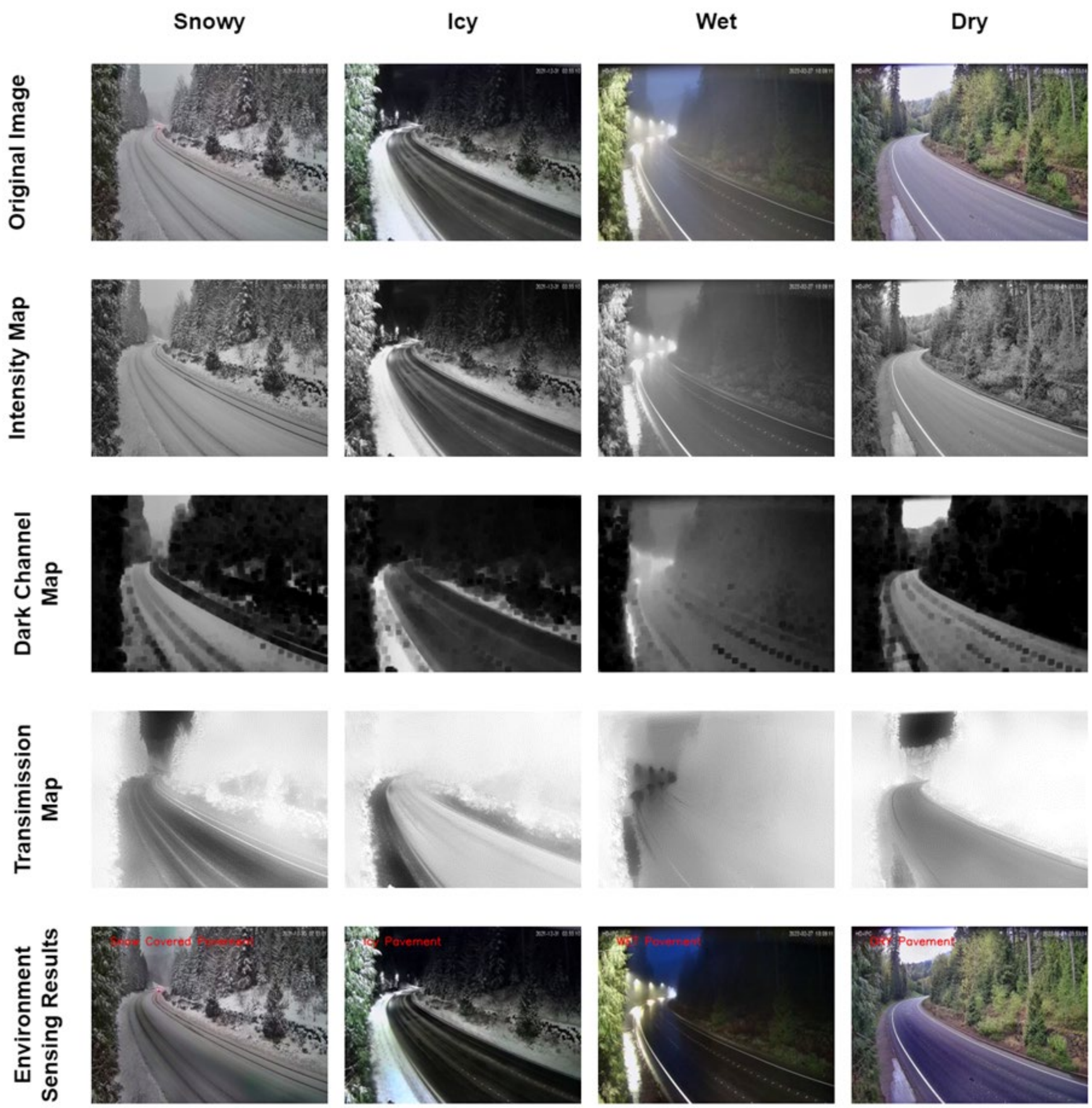


Figure 4-3 Image Processing Results of The Road Surface Conditions Detection (Liu et al. 2023)

Table 4-1 presents the classification results for road surface conditions, achieving accuracy levels of 96% for dry, 92% for wet, 90% for snowy, and 86% for icy surfaces. The higher accuracy for dry and wet conditions is due to their prevalence, providing abundant data for model training. In contrast, snowy and icy conditions are less common in Washington State and primarily occur during the winter, leading to limited data availability and slightly lower accuracy. Additionally, distinguishing between icy and snowy conditions can be challenging during data annotation, potentially causing confusion for the model. Despite these challenges, the overall accuracy of around 95% is satisfactory and supports the practical implementation of a weather condition warning system (Liu et al., 2023).

Table 4-1. Road Surface Condition Test Results (Liu et al. 2023)

Surface Condition	Dry	Wet	Snowy	Icy
Dry	96%	1%	1%	2%
Wet	1%	92%	4%	3%
Snowy	3%	0%	90%	7%
Icy	0%	5%	9%	86%

4.2. Summary

Detecting road surface conditions is crucial for ensuring safe and efficient transportation infrastructure. Technological advancements, including sensors, computer vision, and machine learning, have greatly enhanced the capability to detect and evaluate road conditions. Nevertheless, challenges such as data quality, environmental factors, and cost persist. Future advancements in AI, IoT, and autonomous vehicles are expected to drive further progress in this field, ultimately leading to safer and more reliable roadways.

CHAPTER 5. PEDESTRIAN DETECTION FROM MOUNTED SURVEILLANCE CAMERA TOOL FOR RITI COMMUNITIES

5.1. Pedestrian Detection from Mounted Surveillance Cameras

Pedestrians are a vital component of modern transportation systems, yet they face substantial risks. The latest National Household Transportation Survey (NHTS) indicates that walking accounts for approximately 11% of all recorded trips, making it the second most common mode of transportation (NHTS, 2004). Despite its popularity, walking in urban and densely populated areas presents significant dangers to pedestrian safety. According to the World Health Organization (WHO), pedestrians and bicyclists, collectively known as Vulnerable Road Users (VRUs), make up half of the global fatalities resulting from road crashes (WHO, 2023). Additionally, the Fatality Analysis Reporting System (FARS) shows a steady increase in pedestrian death rates in the United States, rising from 11% in 2002 to 17% in 2020. In 2020 alone, 6,516 pedestrians died in traffic incidents, averaging one death every 81 minutes. The implementation of new technologies holds great promise for reducing severe injuries and fatalities on roads. Advances in sensor and AI technology are crucial in addressing these safety challenges.

Extensive research has been conducted to improve road management, traffic regulation, and traffic monitoring using AI technologies. Cities and agencies are increasingly opting to use existing surveillance infrastructure rather than deploying new sensors, focusing on cost efficiency. This trend is driving the development of intelligent infrastructure, making significant strides in recognizing vehicles, roads, traffic patterns, and VRUs from surveillance camera feeds. As a result, the Intelligent Transportation Systems (ITS) community is increasingly concentrating on pedestrian detection technologies to enhance traffic safety and ensure equity. Pedestrian detection technologies within ITS are generally categorized into three main approaches: detection-based, regression-based, and density estimation techniques.

A significant body of research has focused on enhancing road management, traffic regulation, and monitoring through AI technologies. Cities and agencies are increasingly prioritizing the use of existing surveillance infrastructure over the deployment of new sensors, aiming for cost efficiency. This trend is accelerating the development of intelligent infrastructure, achieving notable progress in recognizing vehicles, roads, traffic patterns, and Vulnerable Road Users (VRUs) from surveillance camera feeds. Consequently, the Intelligent Transportation Systems (ITS) community is emphasizing pedestrian detection technologies to improve traffic safety and ensure equity. These technologies are generally classified into three main categories: detection-based approaches, regression-based approaches, and density estimation techniques.

1. Detection-based Methods: These are key techniques for recognizing objects through feature extraction and classification, playing a crucial role in ITS. Recent advancements have produced sophisticated object detection models such as R-CNN, YOLO, and SSD, which have demonstrated exceptional performance in various applications, including traffic control, forecasting, and management (Zhang, 2021). However, applying these detection-based approaches directly to pedestrian sensing presents challenges due to factors like occlusion, cluttered backgrounds, small object sizes, and motion blur. Unlike larger transportation elements like vehicles, VRUs are smaller and have fewer distinctive features, making them harder to distinguish from their surroundings (Voigt, 2021).

2. Regression-based Methods: These methods leverage the overall characteristics of images, such as texture and gradient features, to identify pedestrians. Global characteristics are analyzed using various regression techniques, including linear regression and Gaussian mixture regression, to detect pedestrians. This approach is particularly adept at handling occlusions, small-sized objects, and blurred images. However, regression-based methods can face limitations due to changes in perspective and the challenge of accurately scaling 2D images, potentially leading to inaccuracies such as overestimations in areas of low pedestrian density and underestimations in densely populated areas.

3. Density Estimation Methods: These methods have been widely used within the computer vision community for analyzing pedestrian traffic in densely populated environments. Unlike detection-based and regression-based techniques, density estimation employs convolutional neural networks (CNNs) to extract complex features and directly estimate crowd sizes using density maps (Liang, 2022). This approach has applications in public safety, the design of public spaces, and the intelligent monitoring of crowds. Despite their potential, incorporating density estimation methods into transportation contexts faces several challenges (Liang, 2022). Ensuring the safety of VRUs requires understanding their precise status and locations, which density estimation methods might not always capture effectively. VRUs tend to cluster in groups of varying densities within transportation environments, and current density estimation techniques often struggle with these variable densities and non-linear scale changes, complicating accurate crowd analysis (Sindagi, 2017).

To achieve accurate and reliable pedestrian detection in transportation scenarios and address safety challenges, the research team proposes an advanced tool incorporating the SARLES algorithm that processes surveillance camera videos of pedestrian crowds frame by frame. This tool can assist traffic agencies in sophisticated transportation management, safety issue monitoring, traffic operation, visual data collection, and more.

5.2. Data Description

The research team employed four datasets to train and evaluate the proposed tool in diverse application settings, including ShanghaiTech, UCF-QNRF, CityStreet, and a custom dataset compiled by the team. This custom dataset was created using live camera feeds from busy intersections and pedestrian zones in Tokyo. These locations were specifically chosen to capture a wide range of challenges, such as complex backgrounds, varied density distributions, and fluctuations in scale and perspective. The dataset comprises 500 annotated images, with 400 designated for training and 100 for testing. The rich variety of urban pedestrian scenarios in this dataset makes it an invaluable resource for refining and validating the tool's performance under real-world conditions.

5.3. Pipeline

Figure 5-1 illustrates the system pipeline, which comprises three main modules: the encoder-decoder module, the density map segmentation and clustering module, and the local patch refinement module.

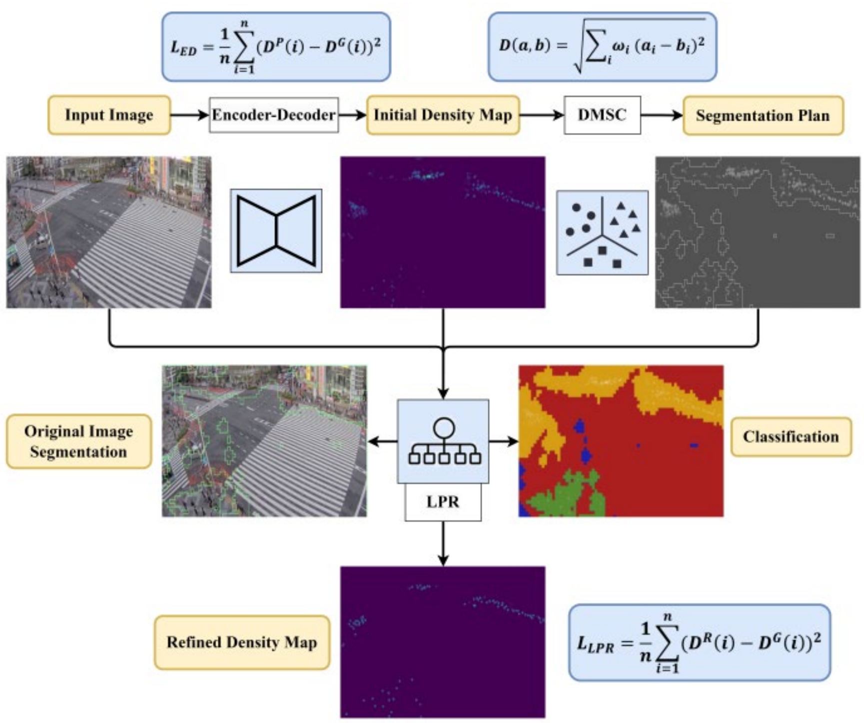


Figure 5-1 Architecture of Pedestrian Detection Framework

The system pipeline consists of three major modules, each playing a critical role in processing and refining pedestrian density maps. The first module, the encoder-decoder module, processes images of pedestrians in traffic scenarios by extracting multi-scale contextual features to generate an initial density map. This map is adequate for basic sensing tasks, such as counting the number of people in a crowd. However, it may lack the detailed information necessary for more advanced tasks, particularly those involving the precise localization of pedestrians within transportation settings.

To address this, the second module, the density map segmentation and clustering module, refines the initial density map by segmenting it into smaller, localized patches based on their density features. This approach is designed to handle the challenges posed by sudden changes in scale and the wide variety of density distributions typically encountered in transportation environments. In these settings, pedestrians often gather in small, distinct clusters across different areas, leading to non-uniform scale changes and diverse density patterns. By segmenting the initial density map into patches that each represent a unique density feature, this module ensures consistency within each segment while maintaining overall diversity. These segmented patches are then processed through the local patch refinement module for further enhancement.

The third module, the local patch refinement module, employs a CNN model with two fully-connected (FC) layers to classify the local patches into five distinct density levels, ranging from high to low. This classification process, combined with the segmentation strategy, provides the necessary input for an ensemble network dedicated to refining the patch density maps. By synthesizing these refined patches, the system can produce an accurate and detailed density map. This enhanced map is instrumental in detecting, sensing, and precisely locating pedestrians within various environments, significantly improving the effectiveness of pedestrian monitoring and safety measures in transportation contexts. Figure 5-2 shows the heatmap of intersection on a raw image generated by the tool.

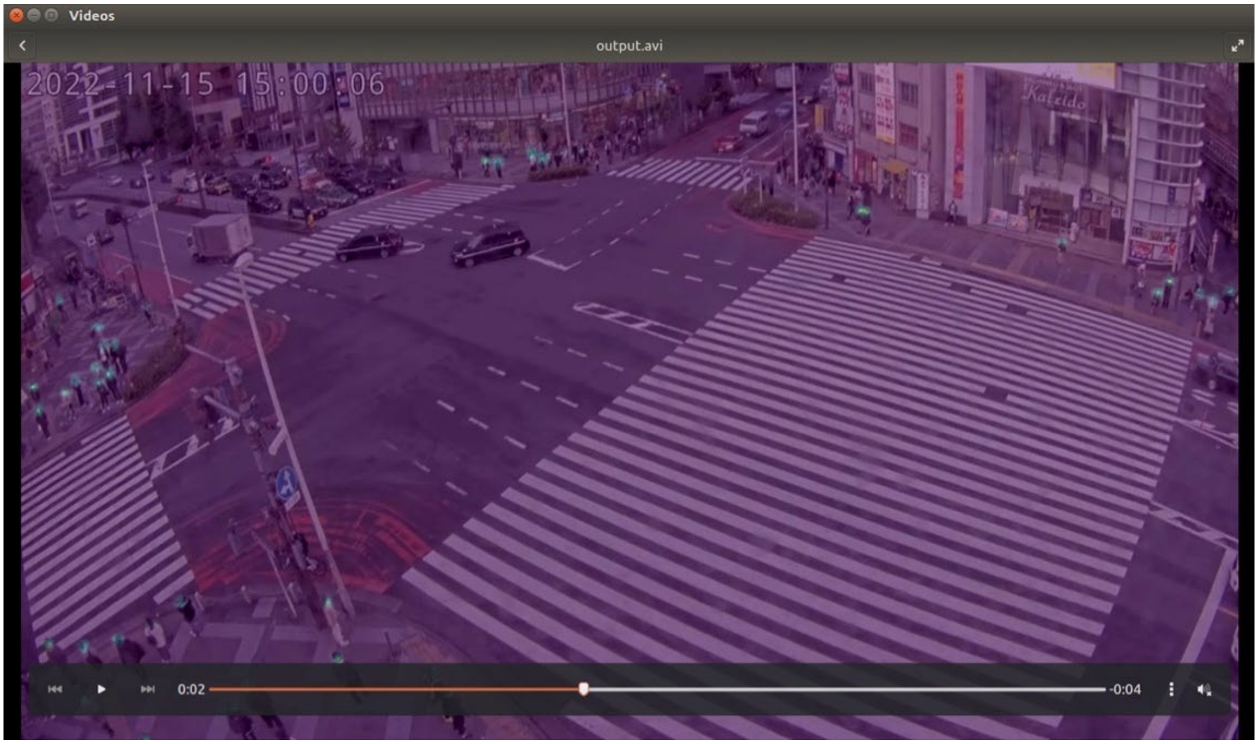


Figure 5-2 Detection result overlaying estimated density heatmap with the raw image.

5.4. Discussion

This tool is an effective approach for identifying and sensing pedestrians within transportation settings, addressing common obstacles to reliable pedestrian detection in complex surroundings. These challenges include intricate occlusions, complicated backgrounds, variations in scale, heterogeneous distributions, shifts in perspective, and the presence of small objects. By adopting a scale-aware strategy, the method effectively captures features across multiple scales and processes representation information. This enhances its capability to detect and count pedestrians across a variety of distributions and complex scenarios, making it a reliable solution for monitoring pedestrian groups in challenging environments.

CHAPTER 6. CLASSIFIED VEHICLE VOLUME TOOL FOR RITI COMMUNITIES

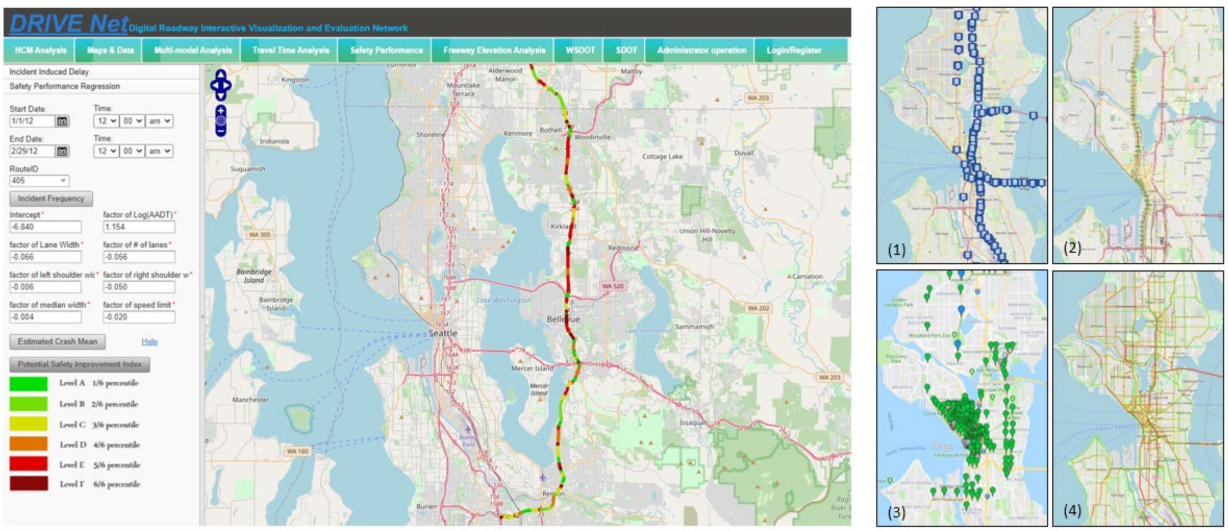
6.1. Background and goal

Classified vehicle volume is a crucial traffic operational metric for transportation agencies. Large vehicles, such as trucks and buses, operate differently from passenger cars due to slower acceleration, inferior braking, and larger turning radii. Adjusting traffic flow parameters for these vehicles is essential for optimizing roadway capacity and enhancing safety. Additionally, the impact of vehicles on pavement deterioration varies significantly by vehicle type, mainly due to differences in axle configurations and total weight. Accurate vehicle classification data allow for precise estimation of pavement wear, leading to more cost-effective pavement design and maintenance strategies. Furthermore, understanding the composition of traffic flow, including the proportion of heavy vehicles, is vital for long-term transportation planning and infrastructure development. It influences decisions on roadway design, bridge structures, and investment in alternative transportation modes.

Loop detectors are widely adopted due to their cost-effectiveness compared to other alternatives. Solutions like Weigh in Motion (WIM) systems and other classification stations can be difficult and costly to maintain due to their complexities, limiting their deployment and coverage of the entire highway system. Mounted traffic cameras face similar limitations. Dual-loop detectors, which use two closely spaced inductive loop sensors embedded in the pavement, detect vehicles based on disturbances in electromagnetic fields caused by the vehicle's metal components. This method facilitates length-based vehicle classification through four steps: vehicle detection, speed measurement, vehicle length measurement, and length-based classification.

Firstly, when a vehicle passes over the first loop, it causes a disturbance in the electromagnetic field. As the vehicle continues, it passes over the second loop, creating another disturbance, marking the vehicle's passage over the dual-loop system. Secondly, the system measures the time taken for the vehicle to travel from the first to the second loop and calculates the speed by dividing the known distance between the loops by the elapsed travel time. Thirdly, based on the estimated speed and occupancy time on each loop, the system calculates the vehicle's length. The length is determined by considering the vehicle's speed and the time interval between the activation of the first loop and the deactivation of the second loop. Lastly, vehicles are classified into categories based on estimated lengths according to predefined length ranges.

In contrast, single loop detectors can measure vehicle count and lane occupancy but cannot measure vehicle lengths, thus are unable to provide classified vehicle volumes. However, single loop detectors are widely available on highway networks and offer an inexpensive alternative to expand classification coverage using existing count stations and traffic operation detectors. Therefore, estimating classified vehicle volumes using single-loop measurements is practically significant. The goal of this tool is to demonstrate how AI and ML can model classified vehicle volume using single loop detector data. This tool aims to extend the capabilities of existing detector stations, allowing urban traffic management systems to better monitor freight traffic within metropolitan areas by leveraging the high density of real-time traffic monitoring stations, as illustrated in Figure 6-1.



(a) safety analysis on DRIVE Net (b) DRIVE Net data sources

Figure 6-1 Demonstration of (a) safety analysis on DRIVE Net; (b) DRIVE Net data sources: (1) loop detector data; blue icons: sensor locations; (2) Verizon cellular data; colored circles: virtual sensor stations; (3) Bluetooth/Wi-Fi based vehicle count data; drops: sensor locations; and (4) INRIX road speed data.

6.2. Data

The loop detector data presented in Table 6-1 was sourced from DRIVE Net (www.uwdrive.net), a platform developed by the University of Washington's STAR Lab. This platform ingests over 50 million lines of loop-detector data daily from the WSDOT sensing system, which has been deployed on Puget Sound area freeways since 2007. Data from two stations were specifically selected for the year 2023.

Table 6-1. Selected loop detector for example study

Station code	Location	Lane No. (from right)*	Dual-loop code	Single-loop code
005es17458	NB I-5 & NE 145th St.	1	_MN__T1	_MN___1
005es18449	SB I-5 & 156th St. SW	1	_MS__T1	_MS___1

* Lane 2 was not included because classified vehicle volume was not reported for the selected year.

6.3. Pipeline

In the WSDOT dual-loop detection system, vehicles are classified into four categories based on their lengths, as shown in Table 6-2: Bin 1 includes vehicles shorter than 20 ft, Bin 2 covers vehicles between 20 ft and 42 ft, Bin 3 includes vehicles ranging from 42 ft to 72 ft, and Bin 4 comprises vehicles longer than 72 ft but shorter than 115 ft. These classifications are applied at loop detector stations, and the data retrieved from the online portal includes the classified vehicle volumes. Depending on the reader's system configuration, the data available for download from these portals may be aggregated in intervals of 20 seconds, 30 seconds, or 5 minutes.

Table 6-2. Length-based vehicle categories used by WSDOT.

Classes	Range of length	Vehicle types
Bin1	Less than 20 ft	Cars, pickups, and short single-unit trucks
Bin2	From 20 ft to 42 ft	Cars and trucks pulling trailers, long single-unit trucks
Bin3	From 42 ft to 72 ft	Combination trucks
Bin4	From 72 ft to 115 ft	Multi-trailer trucks

It's important to note that loop data may have undergone quality control processes, such as imputing missing data for speed records. However, the data still requires pre-processing for our specific purposes. A common outlier in classified vehicle volume data is the reporting of negative values for several or all bins, likely due to sensor malfunctions, which have been addressed by prior quality control steps for speed, occupancy, and volume records.

Once the data is properly pre-processed, the next step is to scope the AI/ML model and its training based on data availability and existing knowledge. The first consideration is which features to include. Since loop detector data can be available every 20 seconds, 30 seconds, or 5 minutes, the chosen model will have varying predictive capabilities. Temporal and spatial correlations are also crucial factors if the agency is interested in understanding how classified volumes correlate over time and with nearby stations. To incorporate these additional features, more complex neural network structures, such as recurrent neural networks and graph neural networks, may be needed.

In this document, the research team demonstrates a basic neural network that ingests features from the single-loop detector and predicts classified vehicle volume at the same time steps. This model does not account for any temporal or spatial relationships. As shown in Figure 6-3, the number of input elements corresponds to the available features: 3 if the loop detector data is reported every 5 minutes (speed, volume, occupancy), 30 if the data is reported every 30 seconds and the user wants to use speed, volume, and occupancy for the last five minutes, 45 if the data is reported every 20 seconds with the same requirements, or 18 if the data is reported every 30 seconds and the user wants to use these features for the last three minutes. The optimal number of hidden nodes is determined through trial-and-error or existing knowledge. The configuration of the output neurons can be 1 if the volumes in the four bins are modeled separately, or 4 if the volumes are modeled together but given different weights when calculating errors. Before training, a train-test split is performed to ensure the model is evaluated on unseen data.

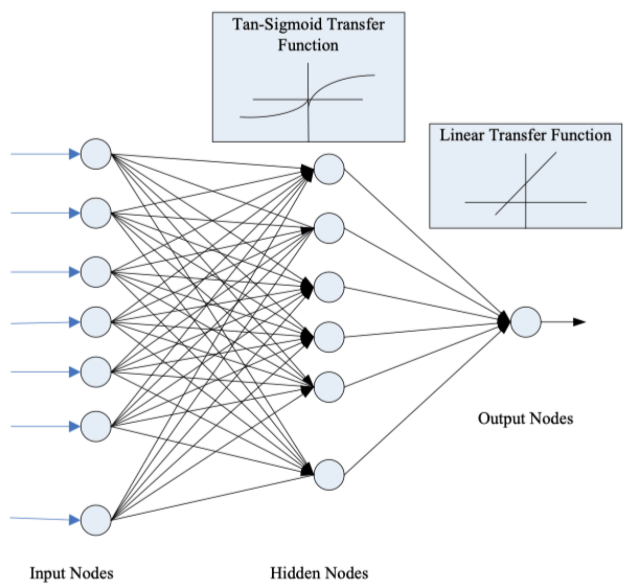


Figure 6-2. Fully connected neural network structure.

6.4. Results

Classified vehicle volumes differ by lane and site. Moreover, volumes in Bin 2 to Bin 4 at the two selected sites exhibit significant daily variation. Consequently, different models are trained for each combination of bin, lane, and site, as shown in Table 6-3.

Table 6-3. Loop detector data aggregated every 5 mins

timestamp	speed	volume	occupancy	Bin1	Bin2	Bin3	Bin4
2023-01-01 00:00:00	60.0	25	2.2	25	1	1	0
2023-01-01 00:05:00	60.0	22	1.8	22	0	0	0
2023-01-01 00:10:00	60.0	17	1.4	20	0	0	0
2023-01-01 00:15:00	60.0	33	2.6	30	0	0	0
2023-01-01 00:20:00	60.0	47	3.6	47	0	0	0

The model shows better performance for Bin 1 volume, as illustrated in Figure 6-4.

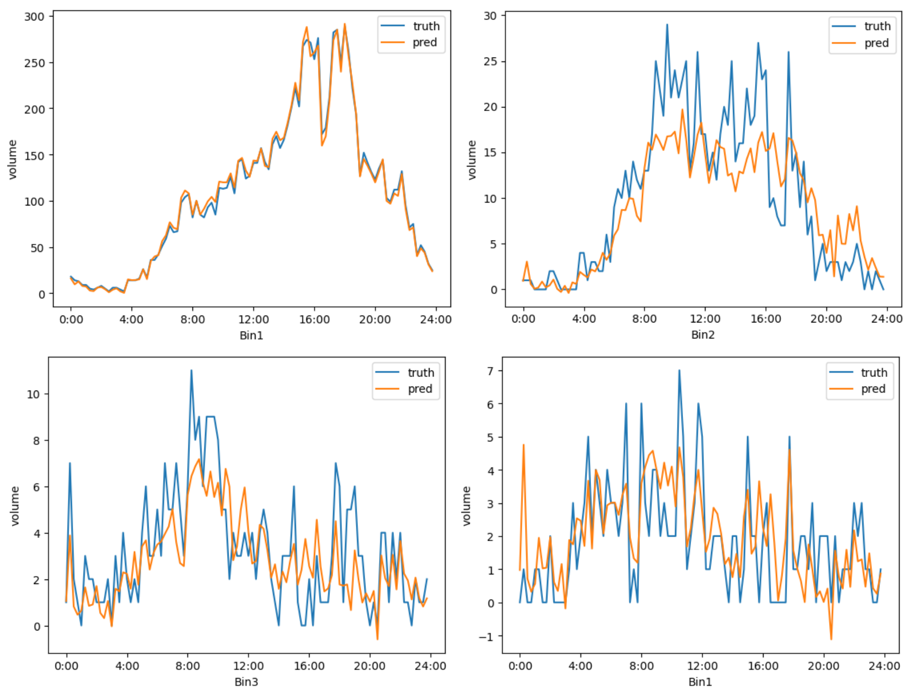


Figure 6-3. Comparison between observed and estimated bin volumes at 15-minute level for detector of 005es17458: _MN___1 on May 23, 2023.

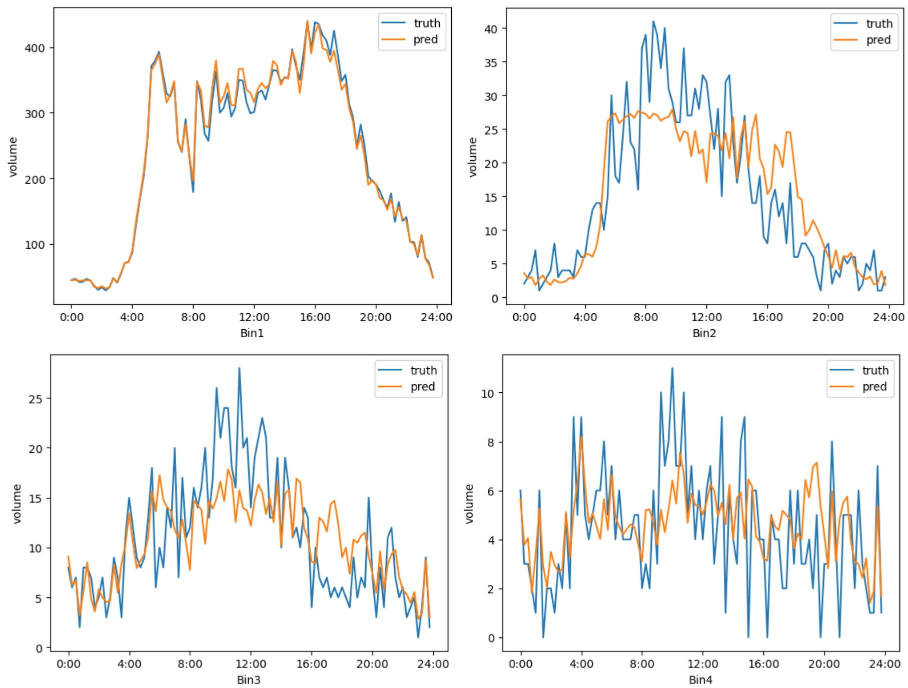


Figure 6-4. Comparison between observed and estimated bin volumes at 15-minute level for detector of 005es18449: _MS___1 on January 11, 2023.

6.5. Discussion

As previously mentioned, several factors affect the model's transferability. Congested conditions can be isolated and analyzed separately, whereas this demonstration considered year-round data, most of which have low to moderate truck volumes. The pre-processing steps included basic quality control but did not incorporate other data sources to validate the aggregate dual loop detector measurements. In other studies, researchers use traffic cameras to further validate classified volumes. Optimal neural network structures may vary by bin, lane, site, and even time frame, as this model is driven by data rather than physics. This variability could limit the model's transferability. Therefore, agencies should understand their operational data and then decide how many models to build and for which sites.

CHAPTER 7. REFERENCES

2001 NHTS User's Guide.** U.S. Department of Transportation, Washington, D.C., 2004.

Federal Highway Administration.** (2012). Traffic Safety Facts "Rural and Urban Comparison". U.S. Department of Transportation.

Heyer, Johanna, Matthew Palm, and Deb Niemeier.** 2020. "Are We Keeping up? Accessibility, Equity and Air Quality in Regional Planning." Journal of Transport Geography 89 (December): 102891. https://doi.org/10.1016/j.jtrangeo.2020.102891.

Jackson, Christopher.** 2006. "Ecoreg Guide." MRC Biostatistics Unit, Cambridge.

Kim, Ellen, Peter Muennig, and Zohn Rosen.** 2017. "Vision Zero: A Toolkit for Road Safety in the Modern Era." Injury Epidemiology 4 (1): 1. https://doi.org/10.1186/s40621-016-0098-z.

Lempert, Robert, James Syme, George Mazur, Debra Knopman, Garett Ballard-Rosa, Kacey Lizon, and Ifeanyi Edochie.** 2020. "Meeting Climate, Mobility, and Equity Goals in Transportation Planning Under Wide-Ranging Scenarios." Journal of the American Planning Association 86 (3): 311–23. https://doi.org/10.1080/01944363.2020.1727766.

Lewis, Elyse O'Callaghan, Don MacKenzie, and Jessica Kaminsky.** 2021. "Exploring Equity: How Equity Norms Have Been Applied Implicitly and Explicitly in Transportation Research and Practice." Transportation Research Interdisciplinary Perspectives 9 (March): 100332. https://doi.org/10.1016/j.trip.2021.100332.

Liang, D., Chen, X., Xu, W., Zhou, Y., & Bai, X.** (2022). TransCrowd: Weakly-supervised crowd counting with transformers. Science China Information Sciences, 65(6). https://doi.org/10.1007/s11432-021-3445-y

Liu, C., Yang, H., Ke, R., Sun, W., Wang, J., & Wang, Y.** 2023. Cooperative and comprehensive multi-task surveillance sensing and interaction system empowered by Edge Artificial Intelligence. Transportation Research Record: Journal of the Transportation Research Board, 2677(9), 652–668. https://doi.org/10.1177/03611981231160174

Martens, Karel, and Aaron Golub.** 2021. "A Fair Distribution of Accessibility: Interpreting Civil Rights Regulations for Regional Transportation Plans." Journal of Planning Education and Research 41 (4): 425–44. https://doi.org/10.1177/0739456X18791014.

Pereira, Rafael, and Alex Karner.** 2021. "Transportation Equity." International Encyclopedia of Transportation, 271–77. https://doi.org/10.1016/B978-0-08-102671-7.10053-3.

Sindagi, V. A., & Patel, V. M.** (2017). A survey of recent advances in CNN-based single image crowd counting and density estimation. Papers With Code. Retrieved from https://paperswithcode.com/paper/a-survey-of-recent-advances-in-cnn-based

StoryMaps by ESRI.** 2021. Yakama Nation. Retrieved from https://storymaps.arcgis.com/stories/4b25c351955e4a4496e5701c9b3df951

Voigt, M., & Cho, H.** (2021). Object Detection: SSD Vs. YOLO. Baeldung on Computer Science. Retrieved from https://www.baeldung.com/cs/ssd-yolo-object-detection

Washington Traffic Safety Commission.** (2019). Washington State Strategic Highway Safety Plan 2019. Washington State Department of Transportation

Wikipedia.** 2024. Yakama Indian Reservation. Retrieved from https://en.wikipedia.org/wiki/Yakama_Indian_Reservation

World Health Organization.** (n.d.). Road traffic injuries. 2023. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

Zhang, X., Liu, Y., & Wang, Z.** (2021). HF-YOLO: Advanced Pedestrian Detection Model with Feature Fusion and Imbalance Resolution. Neural Processing Letters. Retrieved from https://link.springer.com/article/10.1007/s11063-021-10459-z