TOWARD AUTOMATED TRAFFIC SIGN CONDITION EVALUATION UTILIZING VEHICLE-MOUNTED CAMERAS AND LIDAR

A Thesis presented to the Faculty of the Graduate School
at the University of Missouri-Columbia

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Civil Engineering

by
ABIGAIL RENEE WATTS
Dr. Yaw Adu-Gyamfi, Thesis Supervisor
MAY 2025

The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

TOWARD AUTOMATED TRAFFIC SIGN CONDITION EVALUATION UTILIZING VEHICLE-MOUNTED CAMERAS AND LIDAR

Presented by Abigail Renee Watts,

A candidate for the degree of Master of Science in Civil Engineering,

And hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Yaw Adu-Gyamfi

Dr. Praveen Edara

Dr. Clayton Blodgett

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest appreciation to my advisor, Dr. Yaw Adu-Gyamfi, for his support and invaluable guidance throughout my research and writing process. His expertise, patience, and dedication played a crucial role in helping me navigate challenges, develop methodologies, and refine my work. I am truly grateful for his mentorship, which has been instrumental in my academic growth and success.

I am also incredibly thankful for my friends and family (especially my mother Heather, my father Jason, and my sister Chloe) whose unwavering support and encouragement sustained me throughout my graduate studies. Their constant reassurance and belief in my abilities, gave me the strength to persevere and accomplish my goals. I could not have done this without their love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

ABSTRACT

This study proposes a methodology to automate and streamline traffic sign condition assessments and introduces a numerical rating system to evaluate the visual and structural integrity of signs in both rural and urban environments. The goal is to support transportation agencies in identifying damaged, obstructed, or deteriorating traffic signs that may reduce visibility or driver comprehension. The methodology develops a Sign Condition Index (SCI) based on key features, enabling consistent and versatile evaluations. Traffic sign data was collected using an Insta360 ONE X2 camera mounted on a vehicle to capture a video of roadways in and around Columbia, Missouri. Extracted frames were annotated using CVAT (Computer Vision Annotation Tool) to label traffic sign types, and these cropped images were analyzed by a pre-trained large language model (LLM), which generated concise descriptions across four key categories: legibility, color, shape integrity, and surrounding environment. Following manual corrections of the LLM-generated responses, the pairs of cropped images and corrected responses were used to fine-tune the original LLM. This trained version of the model was better equipped to recognize and describe subtle visual features such as fading or physical damage, improving the reliability of automated assessments. In a second data collection phase, a GoPro Hero 11 and a Livox HAP LIDAR sensor were mounted together to capture high-resolution video and dense 3D spatial data. The LIDAR and GoPro datasets were fused and used in model training, enabling automatic detection and classification of signs. Following detection, each sign's condition was scored using the Sign Condition Index. Signs were assigned a value out of four based on yes/no evaluations of the LLM-generated attributes. The process provides a repeatable framework for large-scale traffic sign assessment and maintenance prioritization. A case study utilizing the methods is provided.

CHAPTER 1 INTRODUCTION

1.1 Traffic Signs: An Overview

Cities across the United States are responsible for maintaining an extensive inventory of traffic signs. For instance, Los Altos, California, manages approximately 8,000 traffic signs, while larger municipalities may oversee tens of thousands. Evaluating these signs for compliance with the Manual on Uniform Traffic Control Devices (MUTCD) retroreflectivity standards is a resource-intensive process. The cost of replacing a single sign can range from $100 to $200, not including labor, and prioritizing critical regulatory signs, such as stop or speed limit signs, adds to the expense. The scale and cost of these evaluations and replacements often strain municipal budgets, particularly in smaller cities with limited resources (City of Los Altos, 2011).

Traffic signs are crucial components of road safety and traffic management, serving as visual tools to guide and regulate drivers, pedestrians, and cyclists. The MUTCD sets rigorous standards for sign retroreflectivity, ensuring they remain effective during nighttime or adverse conditions. However, environmental exposure and physical damage often degrade traffic signs, necessitating regular inspections to uphold these standards. **Figure 1. Comparison of Traffic Sign Reflectivity at Night**, illustrates the difference in retroreflectovty levels between various traffic signs. As shown, the signs with improved retroreflective sheeting provide improved nighttime visibility, ensuring compliance with MUTCD standards and enhancing road safety.

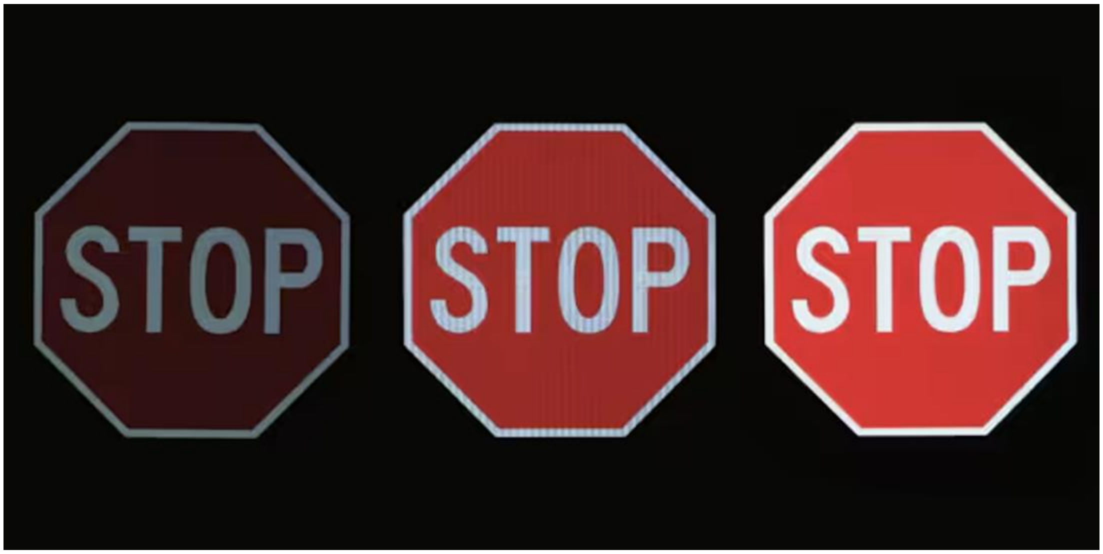


Figure 1. Comparison of Traffic Sign Reflectivity at Night (3M 2009)

Traditional inspection methods, which rely on manual assessment, are resource-intensive and financially burdensome for many municipalities. Smaller cities, in particular, face difficulties due to budget constraints and limited technical capabilities. As Park et al. (2021) highlight in their study, innovative collaborations between municipalities and universities have provided cost-effective solutions by combining academic expertise with practical municipal needs. Their research emphasizes leveraging crash data and prioritizing high-risk signs as effective strategies for improving roadway safety while minimizing inspection costs.

1.2 Problem Statement

Despite the existence of advanced inspection technologies for traffic sign inventorying, their integration is slowed by financial and logistical constraints. Many municipalities, especially smaller ones, continue to rely on inefficient, manual processes to manage traffic sign inventories. This results in signs remaining unchecked for compliance, posing safety risks and increasing liability concerns.

Furthermore, the absence of scalable frameworks that integrate crash data and roadway characteristics into sign prioritization further limits the ability of municipalities to allocate resources effectively. As Park et al. (2021) argue, bridging these gaps requires developing adaptable methodologies that streamline traffic sign inspections and maintenance.

1.3 Objectives of the Study

This study seeks to develop an innovative approach for evaluating traffic signs by simultaneously assessing their retroreflectivity, physical deterioration, and obstructions under real-world conditions. This approach aims to provide municipalities with a practical framework for monitoring traffic sign quality while ensuring compliance with MUTCD standards. To achieve this goal, the study focuses on the following key objective:

- Develop a standardized testing framework that enables daytime assessments of traffic signs while ensuring compliance with MUTCD reflectivity standards. This will be achieved by integrating camera-based visual analysis with LIDAR-derived reflectivity measurements.
- Create a numerical grading system for traffic sign condition, incorporating factors such as color vibrancy, legibility, physical damage, obstructions, and reflectivity performance.

- Leverage a Large Language Model (LLM) to assist in quantifying overall sign condition, using labeled data to train the model to assess sign quality and generate condition scores.
- Develop a versatile methodology that can be applied to both urban and rural environments, making traffic sign monitoring more efficient for municipalities
- Integrate the evaluation process to allow for automated data collection, processing, and condition assessment, improving inventory management and enhancing roadway safety

By combining advanced technologies, this study proposes a simplified and automated process that enhances sign assessment efficiency, reduces subjectivity, and improves overall traffic safety.

Expanded Background Information

1.4 MUTCD Requirements for Traffic Signs

The MUTCD establishes uniform standards for the design, placement, and maintenance of traffic signs. One critical requirement is ensuring retroreflectivity, which allows signs to remain visible under low-light and nighttime conditions by reflecting light from vehicle headlights back to drivers. Retroreflectivity levels must meet minimum thresholds, as specified in MUTCD guidelines, to ensure signs provide adequate visibility for all road users, including those with aging eyesight.

Retroreflectivity is typically measured using handheld devices such as retroreflectometers or through visual nighttime inspection methods outlined in the MUTCD. These evaluations help determine if a sign's retroreflective material has degraded due to environmental exposure, physical damage, or general wear and tear. When retroreflectivity falls below the MUTCD's minimum standards, the sign must be replaced or repaired to maintain compliance (MUTCD). **Figure 2. Retroreflectivity Degradation Overtime** demonstrates the effects of retroreflectivity degradation over time. The left image shows a properly maintained stop sign with sufficient retroreflectivity, while the right image illustrates a faded sign that no longer meets MUTCD standards and requires replacement.

Figure 2. Retroreflectivity Degradation Overtime (FHWA 2018)

Beyond retroreflectivity, the MUTCD also outlines specific requirements for sign legibility placement, and size to ensure visibility and comprehension across various road conditions and environments. For instance, sign lettering must follow the Standard Alphabets for Highway Signs, with uppercase letters typically used except for guide signs, where a mixed case may apply. Letter heigh varies based on roadway type and speed – on urban streets with speed limits of 25 mph (40 km/h), letters must be at least 4 inches (100 mm) high, while rural roads require a minimum of 6 inches (150 mm). For street name signs, lettering should be at least 6 inches (150 mm) tall, and supplementary text should be no smaller than 3 inches (75 mm) (MUTCD, Section 3D). The MUTCD also follows a legibility rule-of-thumb, suggesting 1 inch of letter height per 40 feet (12 meters) of viewing distance.

In terms of placement, the MUTCD requires that signs be mounted at standardized heights and lateral distances to maintain consistency and visibility. Vertical clearance and placement can be influenced by location-specific constraints (e.g., lane width or overhead mounting), but signs must still adhere closely to MUTCD-specified mounting heights and clear zones. Additionally, signs must be positioned to avoid obstruction by vegetation or structures, and spacing guidelines specify consistent interline and border spacing, helping ensure signs are not visually cluttered. Together, these detailed requirements form the basis of proper sign design and maintenance.

In addition to technical standards, the MUTCD outlines five core requirements for effective signs: fulfill a need, command attention, convey a clear meaning, command respect, and allow time for response. Signs should be installed or retained based on engineering judgement, and agencies lacking qualified staff are encouraged to consult their state DOT, LTAP center, or a traffic engineer. Signs fall into three categories – regulatory, warning, and guide – with maintenance urgency varying by type. Critical regulatory signs (e.g., STOP, YIELD) should be fixed within hours, warning signs within three days, and guide signs within about seven working days of discovering deficiency. (FHWA, 2018).

1.5 Traditional Sign Condition Assessment

Historically, sign condition assessments have relied on manual methods, where inspectors physically examine signs for wear, damage, or obstruction. This process is often labor-intensive, requiring trained personnel and significant time investment. For smaller municipalities, the costs associated with these assessments can strain already limited budgets, as highlighted by the "Traffic Sign Maintenance/Management Handbook" by the Minnesota Local Road Research Board (LRRB).

Traditional methods often involve visual inspections conducted from moving vehicles at night to assess retroreflectivity—a process that, according to the Federal Highway Administration (FHWA), relies heavily on subjective judgment and can lead to inconsistencies. These assessments may also involve checklists or handheld retroreflectometers, which are effective but resource intensive. Retroreflectometers, while the most accurate option, cost around $10,000 and are often unaffordable for smaller agencies, though borrowing from a Stat DOT or LTAP may be possible. Both the LRRB and FHWA emphasize that these traditional methods, while necessary for ensuring compliance with nighttime visibility standards, are time-consuming and financially burdensome, particularly for smaller municipalities (FHWA, 2018). **Figure 3 Retroreflectometer Traffic Sign Inspection** below shows the traditional inspection method.

Figure 3. Retroreflectometer Traffic Sign Inspection (FHWA 2018)

1.6 Modern Sign Condition Assessment

Advancements in technology have introduced more efficient methods for evaluating traffic signs. Modern approaches use tools like LIDAR, computer vision, and image processing to automate the inspection process. For instance, vehicle-mounted cameras combined with machine learning algorithms can analyze sign conditions in real-time, identifying issues such as fading, physical damage, or obstruction. Research has demonstrated the effectiveness of integrating LIDAR data with digital images to improve traffic sign detection and recognition, significantly reducing the time and effort required for manual inspections (Guan et al., 2020). **Figure 4. Physically Damaged Traffic Sign,** below illustrates an example of a damaged sign, while **Figure 5. Obstructed Traffic Sign** shows how a sign may appear when obstructed.

Figure 4. Physically Damaged Traffic Sign (Transline Industries 2024)

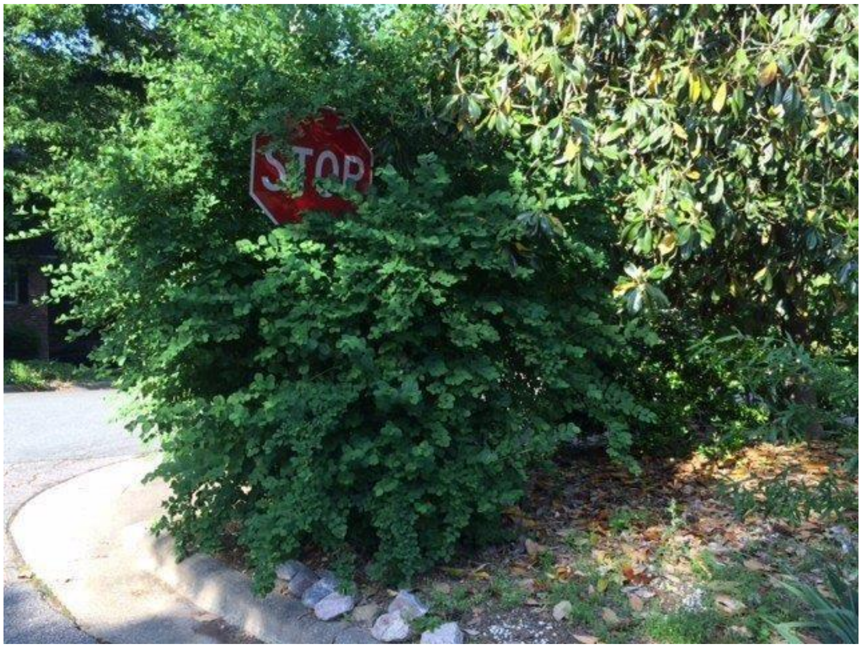


Figure 5. Obstructed Traffic Sign (City of Kennewick 2020)

As shown in **Figures 4 and 5**, damaged and obstructed signs can significantly reduce visibility and driver response time, increasing the risk of accidents. To ensure clear and effective signage, regular condition assessments are essential. Geographic Information Systems (GIS) further support this effort by enabling municipalities to maintain accurate sign inventories, track maintenance history, and prioritize inspections based on factors like crash data and traffic volumes. By integrating GIS, municipalities can streamline operations and make data-driven decisions to enhance roadway safety.

1.7 Structure of the Study

The remainder of the study is organized into multiple chapters. The chapter that follows presents a literature review of existing research on traffic sign condition assessments, highlighting previous approaches and identifying the gaps that this study addresses. The next chapter outlines the methodology, including the data collection process, analysis methods, and the development of the sign condition assessment framework. The subsequent chapter discusses the study's results, covering condition scores and model performance. The final chapter concludes the study with a summary of findings and offers recommendations for future research, addressing the limitations identified throughout the study.

CHAPTER 2 LITERATURE REVIEW

2.1 General Overview

Traffic signs play a crucial role in ensuring road safety by conveying essential information to drivers under various conditions. Retroreflectivity, the property that allows traffic signs to reflect light back towards drivers, is fundamental to their visibility and effectiveness. Retroreflective materials on sign surfaces are designed to reflect headlights, making the signs highly visible even in darkness or adverse weather. Maintaining this reflective quality is essential to avoid hazards related to poor visibility, which can lead to accidents and reduced roadway safety.

Over time, the retroreflective materials on traffic signs deteriorate due to factors such as UV exposure, rain, temperature fluctuations, and traffic-induced wear (Saleh & Fleyeh, 2024). This gradual degradation reduces the brightness and clarity of signs, making them less effective for nighttime visibility. To ensure road safety, transportation authorities must conduct regular assessments to verify that traffic signs meet established retroreflectivity standards. These assessments promote agencies to regularly monitor and replace signs that fall below these thresholds.

Traditional sign assessments often rely on labor-intensive and time-consuming inspections, such as nighttime visual inspections and retroreflectometer measurements, which require assessing each sign individually. While nighttime inspections are subjective and inconsistent, retroreflectometer measurements, though precise, are inefficient and require physical access to each sign (ASTM, 2009). Management strategies like blanket replacement or estimated sign life, as noted by Ai and Tsai (2016), can lead to unnecessary costs or missed critical signs with poor retroreflectivity.

With advancements in technology, new methods, including mobile LIDAR, offer potential for automated and real-time retroreflectivity assessments. These innovations aim to enhance accuracy, efficiency, and consistency in evaluating sign conditions. However, each method faces unique challenges, such as accounting for environmental conditions and physical damage that can impact retroreflectivity. This paper reviews existing research on retroreflectivity assessment and sign condition evaluation, highlighting recent advancements and identifying key limitations that must be addressed to achieve reliable and comprehensive maintenance strategies.

2.2 Sign Detection

The Manual on Uniform Traffic Control Devices (MUTCD) classifies traffic signs into categories to promote consistency and safety across U.S. roadways. These categories include regulatory signs (e.g., stop, yield), warning signs (e.g., sharp curves, school zones), and guide signs (e.g., highway exits). The MUTCD also includes specialized signs like school zone signs, which prioritize high visibility, and temporary signs used in construction zones. These categories are defined by standardized shapes, colors, and symbols to enable quick recognition and response, ensuring a predictable traffic control system (MUTCD, 2009).

In recent years, efforts to improve traffic sign detection have led to the creation of benchmark datasets that support model development for various environments. Zhu et al.

(2016) introduced a benchmark dataset for traffic sign detection and classification. The study consisted of 100,000 images with 30,000 annotated signs from ten regions in China. They utilized a convolutional neural network (CNN) to detect and classify these signs, achieving 84% detection accuracy and 94% recall. However, the dataset had limitations, particularly in its environmental variability, such as sign deterioration and nighttime conditions, which could affect model generalization. These challenges can be addressed by incorporating additional environmental factors in future research to improve real-world applicability.

Another significant contribution to traffic sign recognition is the development of widely used datasets like the German Traffic Sign Recognition Benchmark (GTSRB) and the LISA Traffic Sign Dataset. The GTSRB offers over 50 sign categories under various weather, lighting, and obstructed conditions, making it a valuable resource for training models. Similarly, the LISA dataset includes 47 U.S. sign types with detailed annotations, offering temporal context through video sequences (Møgelmose et al., 2012). While both datasets are valuable for training robust traffic sign detection models, challenges such as overfitting, class imbalances, and the need for effective data augmentation and model fine-tuning remain.

The study by J. Stallkamp et al. (2012) critically evaluates different machine learning algorithms for traffic sign recognition, comparing traditional methods like support vector machines (SVMs) with modern deep learning approaches such as CNNs. The research demonstrates that while classical methods are effective for simpler signs, CNNs excel in complex scenarios involving obstructions, lighting variations, and diverse backgrounds.

The study also benchmarks these models on the GTSRB and LISA datasets, offering insights into their respective strengths and limitations. An important insight from the study is that deep learning models, particularly those using transfer learning and large annotated datasets, achieve superior accuracy but require substantial computational resources. This highlights the ongoing need for optimization techniques to balance model performance with resource constraints in real-world applications.

In a similar vein, Swapna et al. (2021) developed a deep learning-based system that incorporates both detection and classification stages for traffic sign recognition. This system uses HSV color segmentation and shape feature extraction via histogram of oriented gradients (HOG) for the detection phase, and a CNN for improved classification accuracy. Evaluated with the GTSRB dataset, the system demonstrates effectiveness across different conditions, highlighting its potential for integration into advanced driver assistance systems (ADAS). However, challenges remain in recognizing signs under heavy occlusion or severe damage, suggesting areas for further improvement.

2.3 Sign Condition Assessment

Ensuring that traffic signs remain effective over time, assessing both their physical condition and their ability to reflect light for visibility is important. Deterioration in these areas can significantly impact a sign’s functionality and compromise road safety. The following sections examine the methods used to evaluate physical damage, such as wear and environmental impacts, as well as the retroreflectivity of signs, which is critical for traffic sign visibility.

Physical Damage Assessment

Physical damage assessment focuses on identifying signs that are damaged, deteriorated, or otherwise compromised in a way that affects their visibility or legibility. According to the Manual on Uniform Traffic Control Devices (MUTCD), damaged signs should be promptly replaced to ensure proper maintenance and safety standards are upheld. Maintenance activities should account for the sign's position, cleanliness, and visibility during both daytime and nighttime conditions. Additionally, obscured signs resulting from weeds, shrubbery, or other obstructions should be addressed to maintain optimal functionality (MUTCD, 2009, Section 2A.22).

While the MUTCD emphasizes manual maintenance protocols, recent research highlights the potential of advanced technologies to enhance precision and efficiency in sign damage assessment. For instance, Manasreh et al. (2024) developed a framework using mobile LIDAR data to improve predictive modeling of retroreflectivity in road markings. The study extracted over 3,500 features from LIDAR point clouds, capturing fine details of reflective degradation. Although highly accurate, the study’s reliance on controlled conditions limited its ability to generalize findings to real-world scenarios with varying lighting and weather conditions. Additionally, LIDAR’s inability to capture surface dirt or physical damage highlighted a gap in detecting visible deterioration.

Building on these advancements, You et al. (2017) introduced a method integrating mobile laser scanning (MLS), deep learning techniques, and MLS point clouds to identify physical damage in traffic signs. The method effectively identified signs with issues such as tilted poles, deformed boards, and signs that had fallen. It was demonstrated in a real-world scenario following Typhoon Meranti in Xiamen, China, where it successfully updated traffic sign inventories. However, while the approach was highly effective for detecting physical damage, it did not address more subtle forms of deterioration, such as reflective wear or gradual weathering of the sign material. Additionally, the system's reliance on MLS data made it potentially too resource-intensive for routine monitoring, limiting its practical application outside of emergency situations.

Further advancing traffic sign damage detection, Merolla (2024) leveraged computer vision and machine learning models to assess signs exposed to various forms of physical damage, including rust, graffiti, and cracks. Using a custom-built dataset, the study employed a two-step approach: first, detecting signs in images with the YOLOv8 model, and second, classifying their damage state using a convolutional neural network (CNN). This methodology improved the detection of subtle damages over prior approaches. However, challenges like dataset imbalance and the need for retroreflectivity assessment remain unaddressed, leaving room for further refinement.

Retroreflectivity Assessment

Retroreflectivity assessment ensures that signs meet minimum standards for nighttime visibility to support driver safety. The MUTCD mandates that public agencies maintain retroreflectivity levels above specified thresholds, using methods such as visual inspection, retroreflectometer measurements, or scheduled replacements (MUTCD, 20209, Section 2A.08). Retroreflective elements like sheeting types and LED technology are critical for maintaining visibility under low-light conditions (MUTCD, 2009, Sections 2A.07, 2A.08). Retroreflective compliance plays an essential role in ensuring that signs remain legible and effective for road users. To achieve these standards, researchers have developed innovative methods to assess retroreflectivity with greater efficiency.

Ai and Tsai (2016) introduced an automated method for evaluating traffic sign retroreflectivity by using a mobile LIDAR system combined with computer vision techniques. Their model processed LIDAR point clouds to locate traffic signs and then employed image analysis to assess retroreflective quality. This technique significantly enhanced data collection efficiency compared to traditional manual inspections. However, despite its advantages, the model struggled with adapting to differing environmental conditions, such as fluctuating lighting and weather, which affected its consistency. This study demonstrated the growing reliance on data-driven methods in traffic sign assessment while highlighting the challenges in achieving reliable, real-time evaluations under all conditions.

Building on these advancements, Saleh and Fleyeh (2023) conducted a comparative international study of traffic sign retroreflectivity in Sweden and Croatia. They utilized logistic regression, Kaplan-Meier estimators, and artificial neural networks (ANNs) to evaluate each method's effectiveness. The ANN model excelled, achieving up to 94% classification accuracy. Their findings revealed lifespan variations driven by factors such as sign color, location, and sheeting class, emphasizing the need for context-specific maintenance strategies. However, the study's focus on nighttime retroreflectivity left gaps in addressing daytime factors, such as dirt accumulation and UV degradation. Additionally, the absence of a standardized visual scale for classifying sign deterioration highlighted the need for universally applicable metrics to better inform maintenance decisions.

2.4 Infrastructure Condition Degradation Prediction

Early research in infrastructure degradation modeling laid the foundation for predicting pavement marking deterioration. Studies in the 2020s, such as those by Saleh and Fleyeh (2023), focused on understanding how environmental factors, including weather conditions, traffic volume, and material composition, impacted the degradation of road signs and pavement markings. Using regression-based models, they integrated historical environmental and traffic data to estimate retroreflectivity degradation over time. An important innovation in their approach was the inclusion of traffic volume and weather exposure as model inputs. However, these models struggled to account for variability in traffic patterns and differences in sign materials, which limited their predictive accuracy across varied conditions.

Building on earlier research, more recent studies have increasingly focused on predicting retroreflectivity and the service life of traffic signs and pavement markings to support reliable maintenance strategies and ensure road safety. Idris et al. (2024) conducted one of the initial machine learning-based studies on this topic, developing a framework that used random forest models to estimate the degradation of pavement marking retroreflectivity over time. Their model incorporated a diverse range of input data, including traffic load, material type, and environmental conditions, to enhance predictive accuracy. However, while Idis et al. successfully identified key variables, their model was limited by its inability to capture complex interactions among environmental and contextual factors, especially those that might affect signs differently at night versus during the day.

These studies collectively illustrate the evolution of retroreflectivity and durability assessment methods, highlighting both advancements and persistent gaps. While there has been progress in nighttime retroreflectivity testing and precision modeling, no existing study has produced a universally applicable scale that accounts for both reflectivity and physical degradation under varying conditions. Moreover, the absence of daytime-specific assessments limits the applicability of current models, as road signs endure both UV exposure and physical wear, which daytime reflectivity would better capture. Addressing these gaps, particularly through the development of a comprehensive deterioration scale, could support more effective, context-sensitive maintenance strategies.

CHAPTER 3 METHODOLOGY

3.1 Methodology Flow Diagram

A visual overview of the methodology performed in this research can be seen in **Figure 6**. As shown in the figure, two different sensors were used to collect two distinct types of data: a webcam for capturing images of traffic signs and a LIDAR sensor for measuring their reflectivity. Once collected, the data entered a pre-processing phase, which included object recognition through YOLOv8, and traffic sign isolation by cropping them from the original image. This pre-processing procedure was designed to enhance the performance of subsequent visual evaluation by LLM.

The resulting sign image, sign class, and reflectivity value were then input into Qwen 2.5-VL, a multimodal large language model capable of processing both visual and textual data. The model generated descriptive phrases for each visual assessment category—legibility, color, shape integrity, and surrounding environment—and used this information along with reflectivity to calculate an overall condition score.

The final step involved generating a sign condition index (SCI), which provides a comprehensive evaluation of each traffic sign by combining its visual and reflectivity-based assessments, validating the effectiveness of the overall approach.

To evaluate traffic signs condition, both daytime evaluation and nighttime evaluation are important. Based on MUTCD, the retroreflectivity is the only criterion that represents the condition of traffic sign at night. Since the reflectivity intensity value from the LIDAR are positively linear related to the retroreflectivity of traffic sign, the reflectivity intensity value from the LIDAR was used to evaluate the traffic sign condition during nighttime. Hence, the functional section of the proposed framework has been divided into two sections- daytime sign condition evaluation and nighttime sign condition evaluation. For daytime condition evaluation, only images are utilized, and for nighttime condition evaluation, point cloud data associated with image are used.

In the daytime sign condition evaluation process, only images are needed to generate the self-tuned LLM. The pre-trained LLM was asked to give the visual performance from four factors- legibility, color, shape integrity, and surrounding environment through images. Then a self-tuned LLM, which was used to evaluate the traffic sign visual performance, was generated using the corrected annotation with corresponding images and class information. This LLM can generate 3 to 5 descriptive adjectives for four visual performance factors—legibility, color, shape integrity, and surrounding environment—based on a single traffic sign image and its corresponding class label.

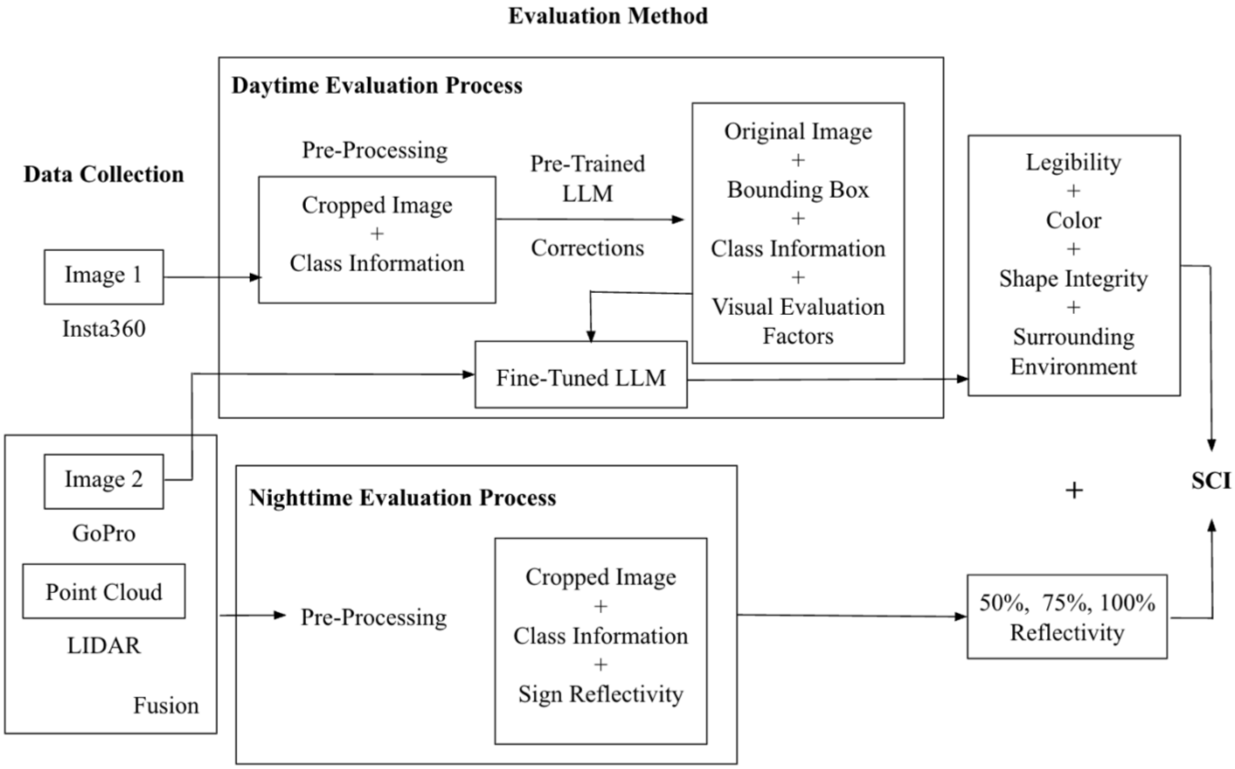


Figure 6. Methodology Flow Diagram

3.2 Methodology Overview

Developing an effective methodology for traffic sign detection and reflectivity evaluation is essential for enhancing road safety and ensuring that drivers can clearly see, read and interpret signs. Traffic signs play a crucial role in guiding drivers to their destinations, enforcing regulations, and preventing accidents. A well-structured method is necessary to produce reliable results that support infrastructure maintenance and safety improvements. The methodology consists of seven key steps: data collection, data preprocessing, data fusion, initial object detection using YOLOv8, visual assessment using a trained LLM, reflectivity analysis, and sign condition scoring. Each step is designed to systematically refine the dataset and optimize model performance for detecting and classifying traffic signs under various conditions.

3.3 Data Collection and Pre-Processing

Data Collection

To develop the dual-mode traffic sign evaluation framework, two rounds of data collection were conducted, each designed to have a distinct component of the evaluation framework-model development and condition inference, respectively. The first round was camera-based and focused on collecting image data to train and test the daytime evaluation model. The second round involved a combined LIDAR-camera setup, designed to generate prediction inputs for both daytime and nighttime evaluations.

The first round of data collection involved recording video footage of traffic signs in and around Columbia, Missouri using an Insta360 ONE X2 camera. The camera was mounted on the hood of the vehicle, as shown in **Figure 7. Initial Data Collection Set Up** and used to capture footage during daytime driving conditions. Although the camera is capable of 360-degree recording, it was positioned to primarily capture traffic signs on one side of the road. Multiple short videos, each around 3 to 5 minutes in length, were recorded to gather a diverse set of traffic signs for model development. This approach provided a manageable and efficient way to collect varied sign imagery while driving through different parts of the study area.

Figure 7. Initial Data Collection Set Up

The second round of data collection utilized a sensor system combining a GoPro Hero 11 Black action camera and a Livox HAP LIDAR Sensor. The GoPro, capable of recording 5.3K video at 30 frames per second, was securely mounted on top of the LIDAR unit. This entire setup was attached to a metal exterior car mount, which was then fastened to the roof of the vehicle. This configuration, shown in **Figure 8. Secondary Data Collection Set Up**, enabled synchronized collection of high-resolution video and detailed 3D spatial data during driving.

The GoPro's wide field of view, enhanced stabilization, and strong low-light performance made it ideal for continuously capturing clear, color-rich footage of traffic signs in varied conditions. Simultaneously, the Livox HAP LIDAR – with its 150-meter detection range and 452,000 points per second at 10 Hz – captured precise depths and reflectivity measurements of the surrounding environment. Together, the two sensors ensured that each traffic sign was recorded with corresponding spatial context, allowing for more advanced analysis of visibility, obstructions, and environmental conditions. Multiple scripts were actively running during collection to monitor system performance and confirm that all equipment was functioning correctly and recording as expected.

Figure 8. Secondary Data Collection Set Up

Reflectivity data was collected using the LIDAR sensor because LLMs, while effective at describing visual features, cannot directly assess a sign’s retroreflective properties from an image alone. Although the model may attempt to infer reflectivity based on brightness or fading, such guesses are unreliable. Recording and uploading the actual reflectivity values for each sign was essential to ensure accurate condition scoring and to produce a reliable Sign Condition Index (SCI).

Reflectivity measurements were recorded during daylight hours, as the retroreflective properties of traffic signs remain consistent regardless of lighting conditions. Because a sign’s reflectivity is an integral material property, it does not fluctuate between daytime and nighttime. Therefore, collecting reflectivity data during the day was both appropriate and efficient, eliminating the need for nighttime measurements while still ensuring an accurate assessment.

Pre-Processing of Image Data

Accurate and rapid object detection is essential for identifying and monitoring specific components within transportation infrastructure. In this study, YOLOv8 (You Only Look Once, version 8), a state-of-the-art object detection model recognized for its high accuracy and real-time processing capabilities, was employed for both object recognition and tracking of traffic signs. A self-tuned YOLOv8 model, specifically customized for this application, was trained on a comprehensive dataset containing over 10,000 manually annotated images representing a wide variety of transportation infrastructure components consistent with the Manual on Uniform Traffic Control Devices (MUTCD). These included stop signs, yield signs, speed limit signs, pedestrian signs, crosswalks, mailboxes, light poles, and utility poles.

The preprocessing workflow began with video and image data, which were passed through the YOLOv8 model to perform object detection and tracking. Each detected sign was assigned a bounding box, a class label, and a unique tracker ID across frames. For every tracker ID, the instance with the highest confidence score or the largest bounding box area was selected to represent the sign. Using this bounding box, the corresponding region of interest (ROI) was cropped from the image, isolating the traffic sign from surrounding visual noise.

This preprocessing step ensured that only the relevant sign region was passed to the visual analysis model, improving downstream large language model (LLM) evaluation by focusing attention on sign-specific features such as color, legibility, and physical integrity without interference from the surrounding environment.

3.4 Dual-Mode Evaluation Framework for Traffic Sign Performance Attributes

3.4.1 Daytime Traffic Sign Condition Evaluation Method

For daytime traffic sign condition evaluation, a large language model (LLM) was trained to generate descriptive assessments of traffic signs based on image visual inputs. To obtain an initial evaluation result, a pre-trained LLM was utilized, then, the evaluation results were corrected. The model takes the original image, specific position of the traffic sign and its class information as input and outputs qualitative evaluations across four visual performance factors: legibility, color, shape integrity, and surrounding environment.

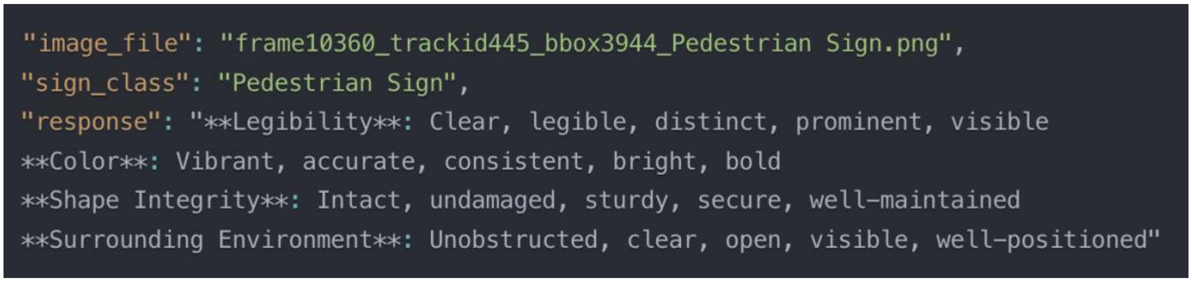
LLM Data Preparation

LLM are advanced AI systems trained on massive datasets to understand and generate human language. Built using deep learning and transformer neural networks, LLMs learn the structure and patterns of language by analyzing large volumes of text. Once trained, they can be adjusted for specific tasks like answering questions, summarizing content, or, in this case, describing features in traffic sign images (Cloudflare, 2025).

Before model evaluation could begin, each pre-processed traffic sign image was cropped to isolate the sign and prepare it for input. These cropped images were then fed into a pre-trained LLM, which was tasked with analyzing signs and generating brief descriptive outputs. Specifically, the model was prompted to describe four key visual categories – legibility (whether the sign's text or symbols were readable), color (vibrancy and signs of fading), shape integrity (whether the sign appeared intact and undamaged), and surrounding environment (if the sign was clearly visible and unobstructed by external factors) – using three to five words for each description.

For example, when shown **Figure 9. Pedestrian Crossing Sign**, from the data collection, the model correctly returned: Legibility – clear, legible, distinct, prominent, visible; Color – vibrant, accurate, consistent, bright, bold; Shape Integrity – intact, undamaged, sturdy, secure, well-maintained; and Surrounding Environment – unobstructed, clear, open, visible, well-positioned. The LLM response for the annotated stop sign is shown in **Figure 10. Pedestrian Crossing Sign LLM Response.**

Figure 9. Pedestrian Crossing Sign

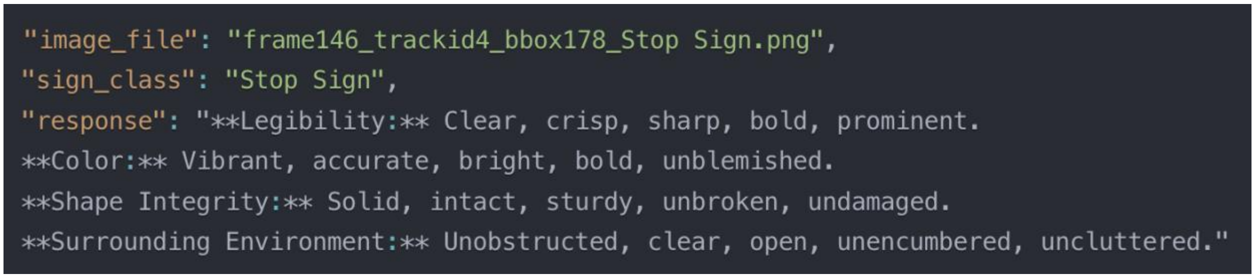


```
"image_file": "frame10360_trackid445_bbox3944_Pedestrian Sign.png",
"sign_class": "Pedestrian Sign",
"response": "**Legibility**: Clear, legible, distinct, prominent, visible
**Color**: Vibrant, accurate, consistent, bright, bold
**Shape Integrity**: Intact, undamaged, sturdy, secure, well-maintained
**Surrounding Environment**: Unobstructed, clear, open, visible, well-positioned"
```

Figure 10. Pedestrian Crossing Sign LLM Response

Following this automated generation, each model response was reviewed for accuracy to ensure the descriptions matched the actual visual characteristics of the traffic sign. This manual verification step helped maintain labeling quality and minimized the risk of inconsistencies or misinterpretations, which could otherwise affect model performance in later stages. For instance, this accuracy-checking process was applied to signs like the one shown in **Figure 11. Faded Stop Sign**. In this example, the stop sign is clearly faded and illegible, but the LLM-generated response – shown in **Figure 12. Faded Stop Sign LLM Response** – failed to accurately capture its poor condition. This highlighted occasional limitations in the LLM's visual interpretation, especially when signs were severely deteriorated.

Figure 11. Faded Stop Sign



```
"image_file": "frame146_trackid4_bbox178_Stop Sign.png",
"sign_class": "Stop Sign",
"response": "**Legibility:** Clear, crisp, sharp, bold, prominent.
**Color:** Vibrant, accurate, bright, bold, unblemished.
**Shape Integrity:** Solid, intact, sturdy, unbroken, undamaged.
**Surrounding Environment:** Unobstructed, clear, open, unencumbered, uncluttered."
```

Figure 12. Faded Stop Sign LLM Response

The whole dataset was divided into training and test sets by a portion of 70% and 30%.

Fine-tuned LLM Generation

After the correction step, the original image, the position and class of each traffic sign, along with the corresponding visual condition descriptions, were compiled as training data for the model. A self-tuned large language model was developed based on Qwen-VL 2.5, a vision-language model capable of understanding and generating multimodal outputs.

The model was fine-tuned using many distinct traffic sign examples. These inputs were formatted into structured prompts to guide the model in producing 3–5 descriptive adjectives for each of four visual performance factors: legibility, color, shape integrity, and surrounding environment. The output format was carefully designed to maintain consistency and interpretability for downstream evaluation tasks.

Supervised fine-tuning was conducted using the corrected descriptors as reference outputs. The dataset included a variety of sign types and conditions to ensure model robustness and generalization to diverse real-world scenarios.

3.4.2 Nighttime Traffic Sign Condition Evaluation Method

For nighttime traffic sign condition evaluation, reflectivity was assessed using intensity values extracted from LIDAR point cloud data. Based on the Livox HAP LIDAR sensor, reflectivity intensity values range from 0 to 255. According to standard thresholds, signs with intensity values below 150 are considered to have insufficient or non-compliant reflective properties for traffic signs.

Camera and LIDAR Data Fusion

To achieve effective fusion of LIDAR and camera data, both spatial and temporal calibration were conducted. The system operated using the Robot Operating System

(ROS), ensuring precise synchronized data acquisition through timestamping. Spatial calibration began with intrinsic calibration, which corrected image distortion and established an accurate 3D-2D projection relationship. Then, extrinsic calibration was performed by aligning corresponding points from the LIDAR point cloud and camera images, optimizing the rotation and translation matrices to transform between the two coordinate systems.

With these parameters, the 3D LIDAR point cloud was mapped on to the 2D image plane, integrating depth and intensity data from LIDAR with the texture and color information from the camera. This fusion process allowed for a comprehensive analysis of traffic signs, combining visual attributes such as text clarity, color vibrancy, and physical condition with LIDAR-derived data. By merging these datasets, the model gained a more robust understanding of sign visibility and deterioration, enhancing its ability to accurately assess real-world signage conditions.

Figure 13. Before Data Fusion and **Figure 14. After Data Fusion** illustrate the effectiveness of the fusion process. These images highlight the contrast in reflectivity between the traffic sign and its surrounding environment, demonstrating the accuracy of the fusion. Reflectivity values in the point cloud were color-coded to enhance visual assessment, with each sign containing multiple reflectivity points. From these, both the 50th and 75th percentile reflectivity values were extracted and analyzed.

Figure 13. Before Data Fusion

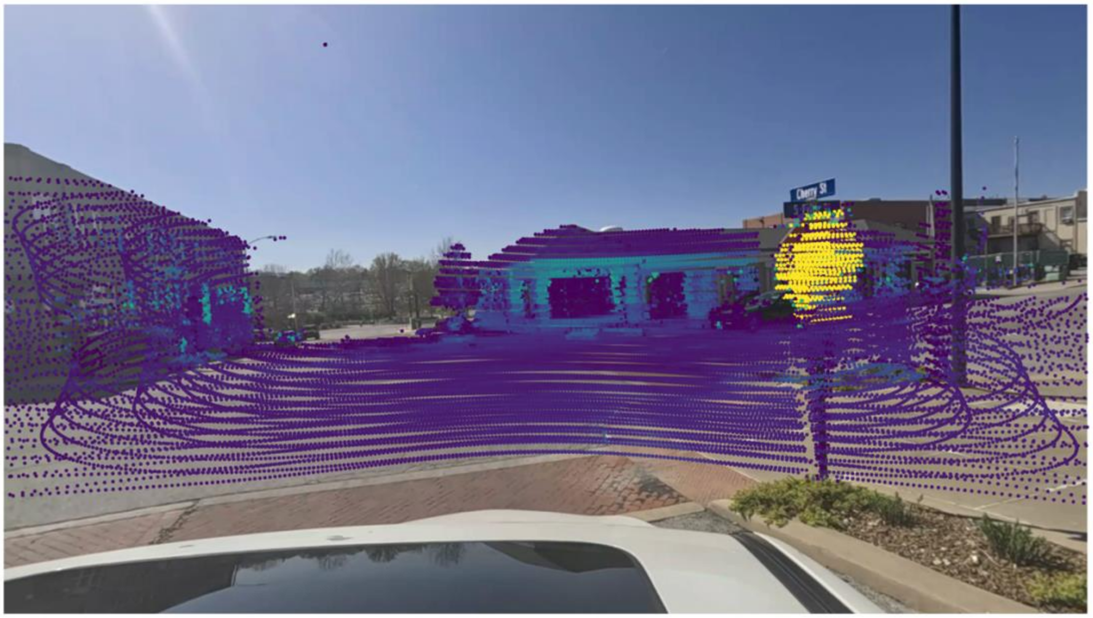


Figure 14. After Data Fusion

Reflectivity was measured on 0-255 scale. Values between 255-221 were assigned a yellow color, while values around 150 were assigned a green color, and values below 150 were assigned purple color. This allowed for a quick, visual evaluation of reflectivity levels. **Figure 15. High Reflectivity Traffic Sign** shows a traffic sign with an overall reflectivity of 255, whereas **Figure 16. Low Reflectivity Traffic Sign** displays a sign with an overall reflectivity of 164, demonstrating the range of reflectivity captured in the dataset.

Figure 15. High Reflectivity Traffic Sign

Figure 16. Low Reflectivity Traffic Sign

3.5 Sign Condition Index (SCI) Generation

After completing the dual-mode traffic sign evaluation method, the next step was to develop a scoring system to quantify the overall sign condition. This led to the creation of the Sign Condition Index (SCI) - a flexible 10-point scoring system that combines visual analysis from a fine-tuned LLM with LIDAR-derived reflectivity data.

A technique known as prompt engineering was used to guide the model's evaluation based on the four key attributes that were extracted from the previous evaluation method-color, legibility, physical integrity, and surrounding environment. Prompt engineering is the practice of developing specific and effective inputs to guide AI models toward desired outputs. This technique allows developers to fine-tune outputs, mitigate bias, and guide the model to give accurate results from LLMs (Schmitt, 2025). By refining prompts with clear instructions and context about traffic sign assessment, the LLM was able to produce descriptive evaluations and assign an initial 0-5 score for each attribute. In addition, the reflectivity factor was also scored on a scale from 0 to 5 based on the 50th, 75th, and 100th percentile reflectivity values of each sign.

To ensure that both daytime and nighttime performance were meaningfully represented, the SCI was computed using a weighted combination of the five individual scores. Specifically, the four daytime evaluation visual factors were aggregated and scaled by a weight of 0.3, while the reflectivity score was weighted more heavily at 0.7, reflecting the critical role of nighttime visibility in traffic safety. A constant of 0.5 was added to serve as a baseline, ensuring that even severely deteriorated signs retained a minimum nonzero score. The final formula is shown in **Equation 1. SCI**.

Equation 1. SCI

$$SCI = 0.3 \times (Legibility + Color + Shape\ Integrity + Surrounding\ Environment)$$

$$prompt\ score + 0.7 \times Reflectivity\ Score + 0.5$$

The proposed SCI provides a comprehensive and balanced measure of traffic sign condition by incorporating daytime and nighttime performance. It addresses the imbalance in factor quantity through weighted aggregation and reflects the relative importance of nighttime visibility, giving a fair and interpretable scoring system.

While this study placed greater weight on reflectivity to account for nighttime driving safety, the SCI scoring system remains highly customizable. Municipalities can assign different weights to prioritize specific aspects based on their local needs—for instance, placing more importance on the surrounding environment in areas with high levels of visual clutter, such as urban corridors with dense signage, vegetation, or commercial advertisements that may obstruct or compete with traffic signs.

CHAPTER 4 RESULTS AND DISCUSSION

4.1 Fine-Tuned Daytime Traffic Sign Condition Evaluation Model

Evaluation Metrics

To evaluate the quality of the descriptive adjectives generated by the fine-tuned LLM for each visual performance factor, a semantic similarity–based precision and recall framework was adopted. For every traffic sign instance, the LLM predicted adjectives are compared against the human-annotated ground truth descriptors. This evaluation captures both the semantic relevance of the predicted outputs and the extent to which they cover the key concepts identified in the human-annotated ground truth.

Precision is calculated by comparing each predicted word to all ground truth descriptors and recording the maximum semantic similarity for each. This measures how closely the model-generated adjectives align with the reference descriptions, indicating the relevance and appropriateness of the output.

Recall, on the other hand, is calculated by comparing each ground truth word to all predicted descriptors and identifying the maximum semantic similarity for each. This assesses whether the model successfully included the essential aspects reflected in the ground truth annotations.

Semantic similarity is computed using cosine similarity, and scores are averaged across all samples. By evaluating in both directions, a final score that averaging the precision and recall was applied to provide a balanced measure of the model's descriptive performance in terms of both accuracy and completeness. The final formula is shown in **Equation 2. Final Score Averaging.**

Equation 2. Final Score Averaging

$$precision = \frac{1}{n} \sum\nolimits_{\omega \in PRED} \ max\cos sim_{(\omega,\, g)}$$

$$recall = \frac{1}{m} \sum\nolimits_{\omega \in GT} \ max\cos sim_{(g,\omega)}$$

$$f1 = \frac{(precision + recall)}{2}$$

Where n and m denote the number of values in prediction and ground truth sets respectively. w and g are the values from prediction and ground truth.

Evaluation Results

In the test set, a total of 224 distinct traffic sign images of 36 classes were used to evaluate the performance of the fine-tuned Daytime Traffic Sign Condition Evaluation Model. The distribution of $f1$ scores across four visual performance factors is presented in **Figure 17. Visual Performance $f1$ Scores**, and the overall $f1$ scores are shown in **Figure 18. Overall $f1$ Scores.** Additionally, **Table 1. Mean and Median $f1$ Scores** summarizes the mean and median $f1$ score across the test set.

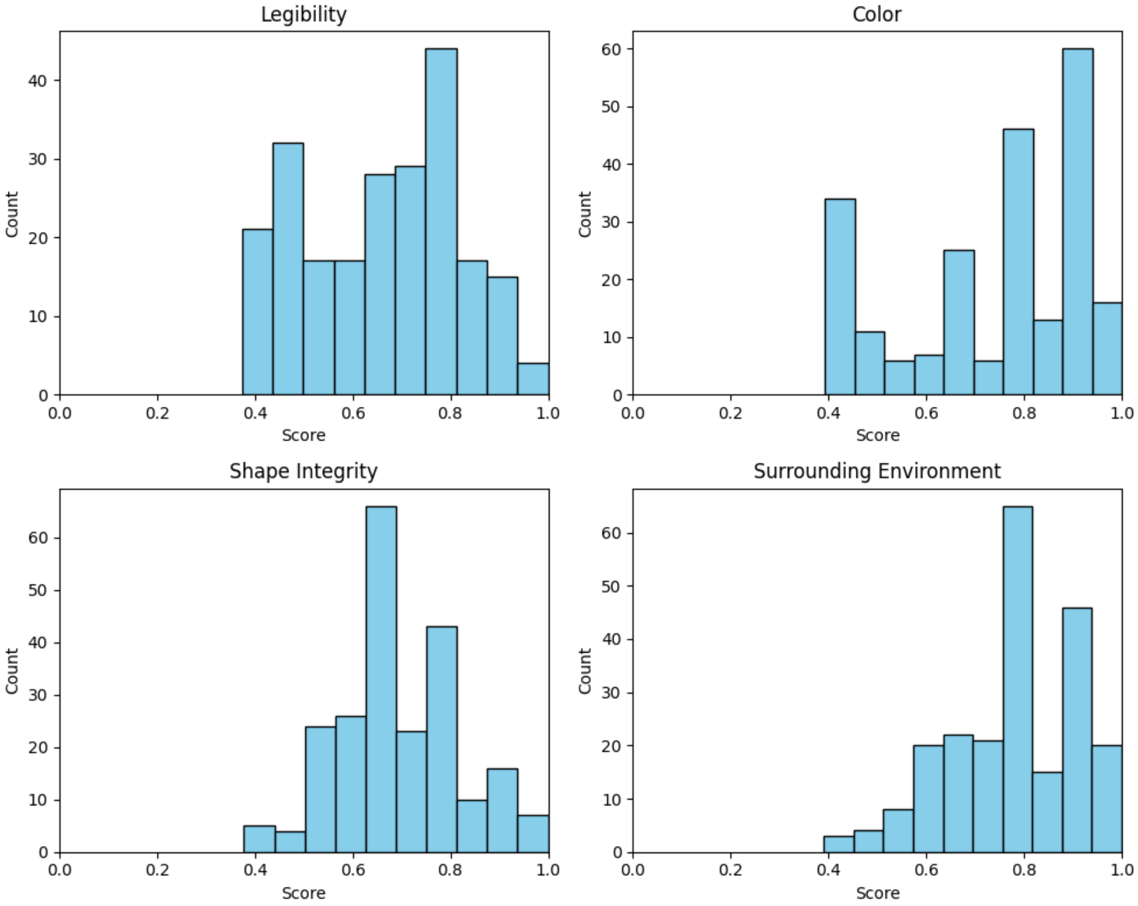


Figure 17. Visual Performance $f1$ Scores

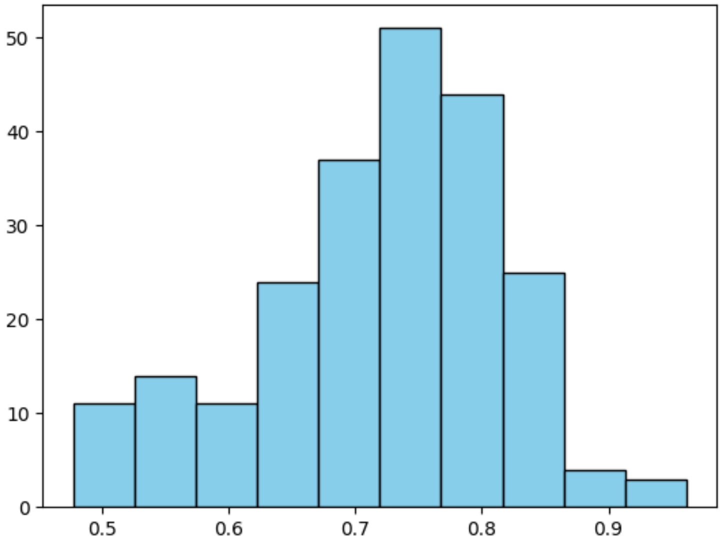


Figure 18. Overall $f1$ Scores

Table 1. Mean and Median $f1$ Scores

Statistics	Legibility	Color	Shape Integrity	Surrounding Environment	Overall
Mean	0.660	0.738	0.694	0.782	0.718
Median	0.681	0.789	0.679	0.792	0.732
Std	0.157	0.185	0.124	0.132	0.098

Based on the results shown in **Figure 18** and **Table 1**, the factor surrounding environment exhibited the highest performance with a mean $f1$ score of 0.782, and a median of 0.792, indicating a consistent predictions with relatively low variability (std = 0.132). Factor color also archives strong performance with a mean value of 0.738 and a median value of 0.789. However, the standard deviation value of 0.185 suggests more variation in prediction result. This may be attributed to the sensitivity of visual inconsistencies, such as lighting condition and resolution of images.

Factor shape integrity exhibited a stable performance with a mean value of 0.694, median value of 0.679, and a standard deviation of 0.124, which reflect the reliability in identifying deformation and distortion in shapes. In contrast, the legibility showed the lowest mean and median score of 0.660 and 0.681, indicating the inconsistency in evaluating the text and pattern clarity. This is potentially caused by the different variability in sign class, including various font size and symbols, or occlusion issue. Finally, the overall score showed a relatively strong central tendency with low variability, demonstrating the generalization capability and robustness of the model.

4.2 Dual-Mode Traffic Sign Evaluation Framework

After the development of the traffic sign evaluation and SCI generation framework, a new dataset was used to assess its effectiveness. This dataset consists of 190 unique traffic signs in 33 classes, all collected in the second round of data. Each sample includes both a traffic sign image and its corresponding reflectivity intensity values obtained from LiDAR measurements.

Following the established dual-mode traffic sign condition evaluation framework, the images were fed into the daytime evaluation module, which utilized the fine-tuned LLM to generate descriptive outputs across four visual performance factors: legibility, color, shape integrity, and surrounding environment. Simultaneously, the LiDAR- based reflectivity values were processed using the nighttime evaluation module, which computes the 50th ,75th, and 100th percentile reflectivity values for each sign.

The results from both components were then integrated through the third stage of the framework—SCI generation. Using a prompt-based scoring mechanism, each factor for every sign was assigned a score ranging from 0 to 5, resulting in a total score of 10. According to the proposed SCI formulation, nighttime performance was assigned a higher weight of 0.7, while the daytime visual factors were weighted at 0.3. Based on this dual-mode traffic sign condition evaluation framework, the final SCI for the 190 evaluated signs ranged from 5.3 to 10 with a mean of 9.31 and standard deviation of 1.139. The distribution of final SCI is shown in F**igure 19. Distribution of Final SCI**

Values. From this figure, most signs received high ratings, with over 50% achieving the maximum score, indicating generally good overall condition.

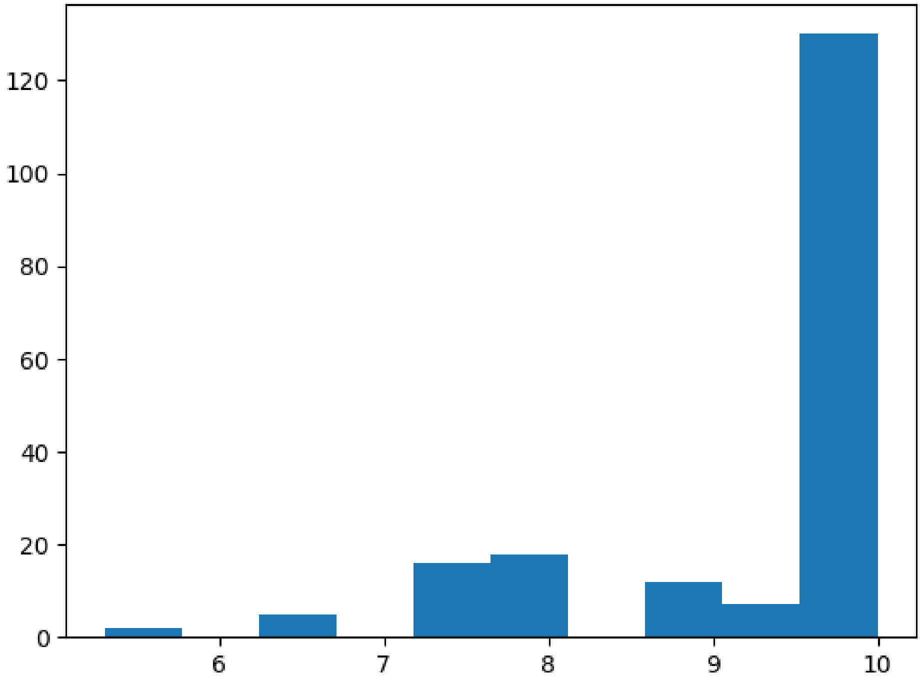


Figure 19. Distribution of Final SCI Values

When the model was given a regulatory sign, **Figure 20. High-Quality Traffic Sign**, in excellent condition with a high reflectivity, approximately 255, it produced a high SCI score of 10 out of 10. In contrast, a different sign in poor condition, the Bike Lane Only Sign in **Figure 21. Low-Quality Traffic Sign**, with faded appearance and low reflectivity, approximately 152, 153, 159 at 50th, 75th, and 100th percentile, received a significantly lower SCI score of 6.5 out of 10. This contrast highlights the model's ability to accurately assess a sign across a range of conditions using both visual and reflectivity inputs.

Figure 20. High-Quality Traffic Sign

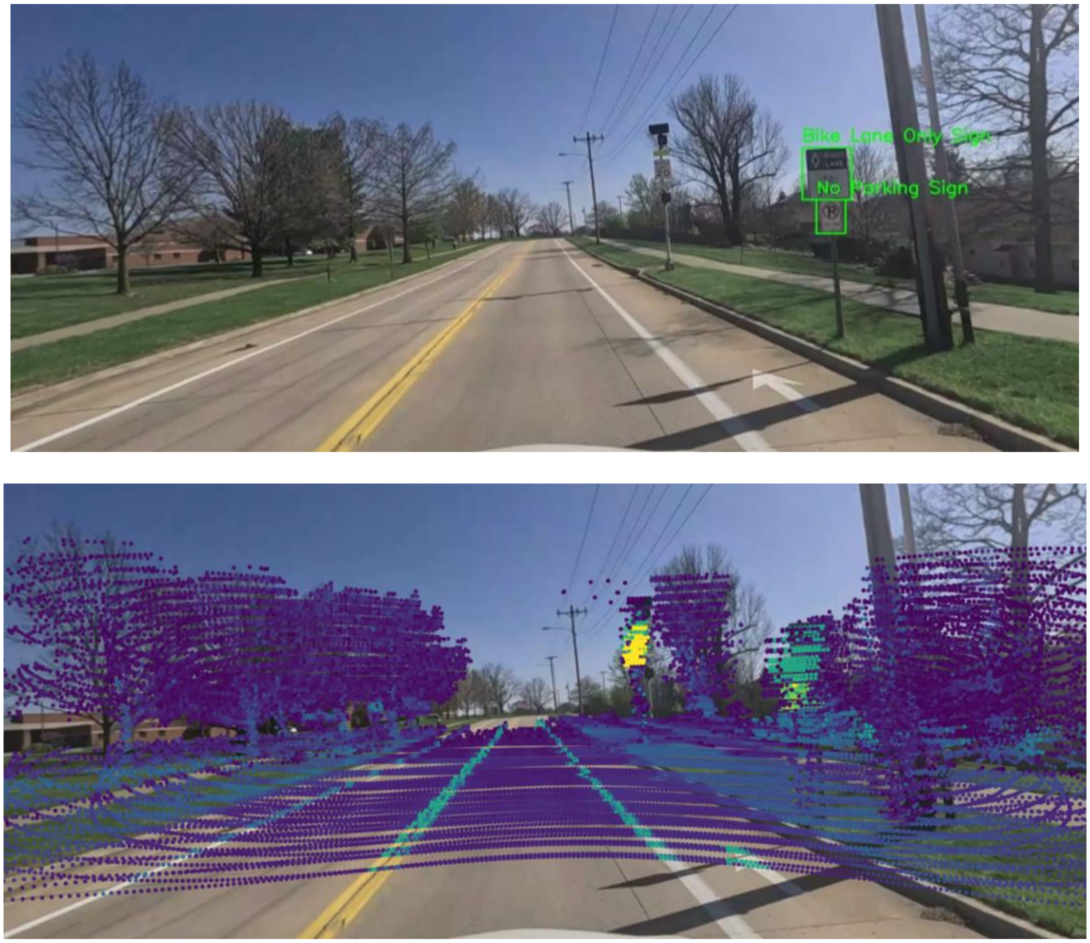


Figure 21. Low-Quality Traffic Sign

CHAPTER 5 SUMMARY

In conclusion, this study presents a novel and cost-effective approach to the automated evaluation of traffic sign conditions. The methodology combines data from a GoPro Hero 11 Black action camera and a Livox HAP LIDAR sensor, which were mounted on a vehicle to collect both visual and spatial data. The images captured by the GoPro were annotated and analyzed using a pre-trained LLM to evaluate the signs' legibility, color, shape integrity, and surrounding environment. The LLM-generated data was then combined with the LIDAR data to classify and detect traffic signs. The refined responses from the trained LLM were then combined with the LIDAR data to classify and detect traffic signs.

A Sign Condition Index (SCI) was developed to evaluate the physical appearance of each sign as well as its reflectivity, using the LIDAR sensor's readings. The model demonstrated strong performance in detecting and classifying traffic signs, with a reliable process that allows for efficient evaluation of traffic sign conditions. The SCI, which is based on the sign's score for physical condition and reflectivity, provided a comprehensive assessment of each sign's usability. Despite its effectiveness, potential limitations include the lack of data collection during periods of inclement weather, limited availability of rare or severely deteriorated signs, and variability in vehicle speed during data collection. These factors may influence the consistency of sign detection and condition evaluation.

The methodology, while innovative, highlights areas for further refinement, such as ensuring consistent data collection under all environmental conditions and improving the handling of speed-related inconsistencies. Despite these limitations, the proposed system has the potential to significantly enhance the efficiency of traffic sign inspections. By automating and simplifying the process, municipalities can better prioritize maintenance, improving road safety and infrastructure management.

CHAPTER 6 CASE STUDY

To demonstrate the practical application of the proposed methodology, an example is provided. This example offers insight into how the methods can be applied in a real-world scenario to automate the assessment of traffic sign conditions using both visual and reflectivity-based data.

The initial data collection occurred over multiple roadways in Columbia, Missouri, on a sunny day. Multiple 3-5-minute-long video segments were recorded using the Insta360 ONE X2, a 5.7k 360° action camera capable of capturing immersive, high-resolution footage at 30 frames per second. The camera was mounted on the hood of a car, and the vehicle followed typical urban and residential routes with speed limits ranging from 25-40 mph. The study area for this phase is shown in **Figure 22. Example Initial Study Area**. While the figure does not depict all routes captured during the phase, it highlights seven representative paths used in the data collection, each displayed in a distinct color to differentiate the routes.

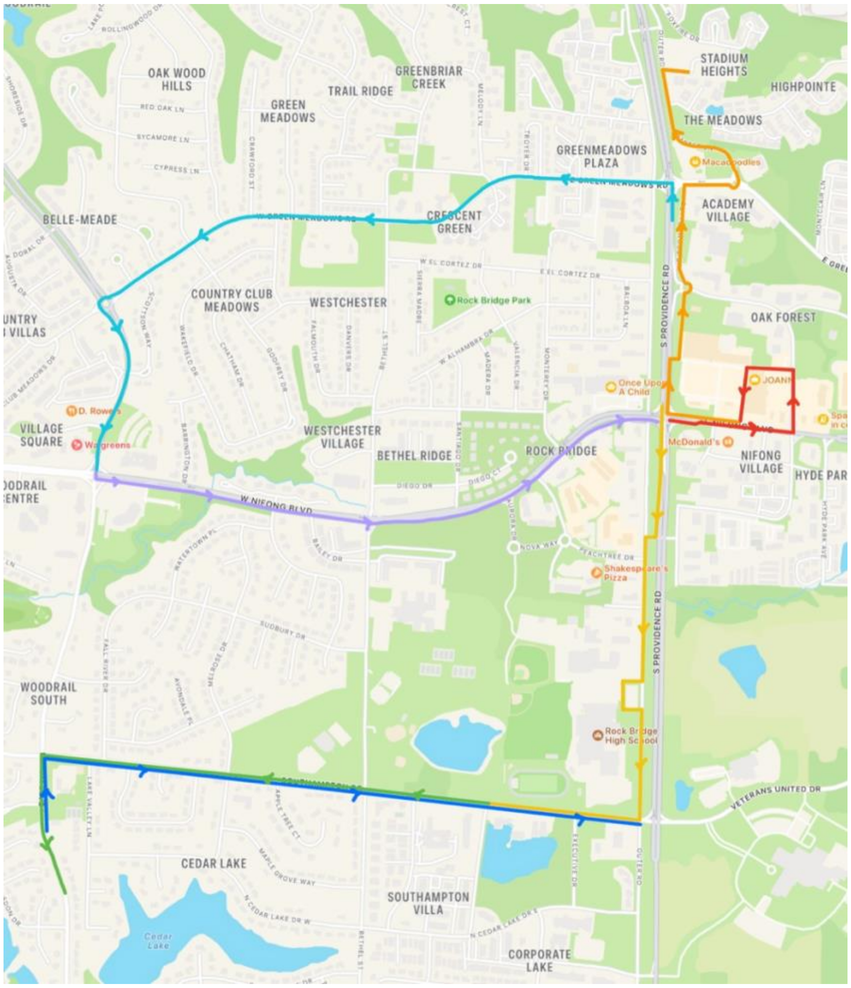


Figure 22. Example Initial Study Area

In post-processing, video frames were extracted and uploaded to CVAT, where traffic signs were labeled according to their type (e.g., stop, yield, speed limit). Once labeled, individual sign instances were cropped into static images. These cropped images were input into a pre-trained large language model (LLM), which was prompted to generate four to five descriptive words for each of the following visual condition categories: legibility, color, shape integrity, and surrounding environment. The model's responses were reviewed and corrected as needed to ensure accuracy. The corrected outputs, along with the cropped images, were then used to fine-tune the LLM for visual condition assessment. This phase focused on visual evaluation and laid the foundation for the sign condition assessment framework.

A second round of data collected was conducted over different roadway segments in Columbia, Missouri, also on a sunny day. This time, video and reflectivity data were collected using a GoPro Hero 11 Black and a Livox HAP LIDAR sensor. The GoPro Hero 11, mounted to the vehicle's roof, recorded a 5.3K video at 30 frames per second, capturing high-clarity images of roadside signage. The Livox HAP, a high-precision, long-range LIDAR sensor with a 300-meter detection range, recorded reflectivity data simultaneously.

The areas covered during this second phase are shown in **Figure 23. Example Secondary Study Area**. After data collection, camera and LIDAR outputs were synchronized and fused through preprocessing steps that included timestamp alignment, calibration, and coordinate transformation. This fusion ensured that each visual frame was accurately paired with its corresponding reflectivity values.

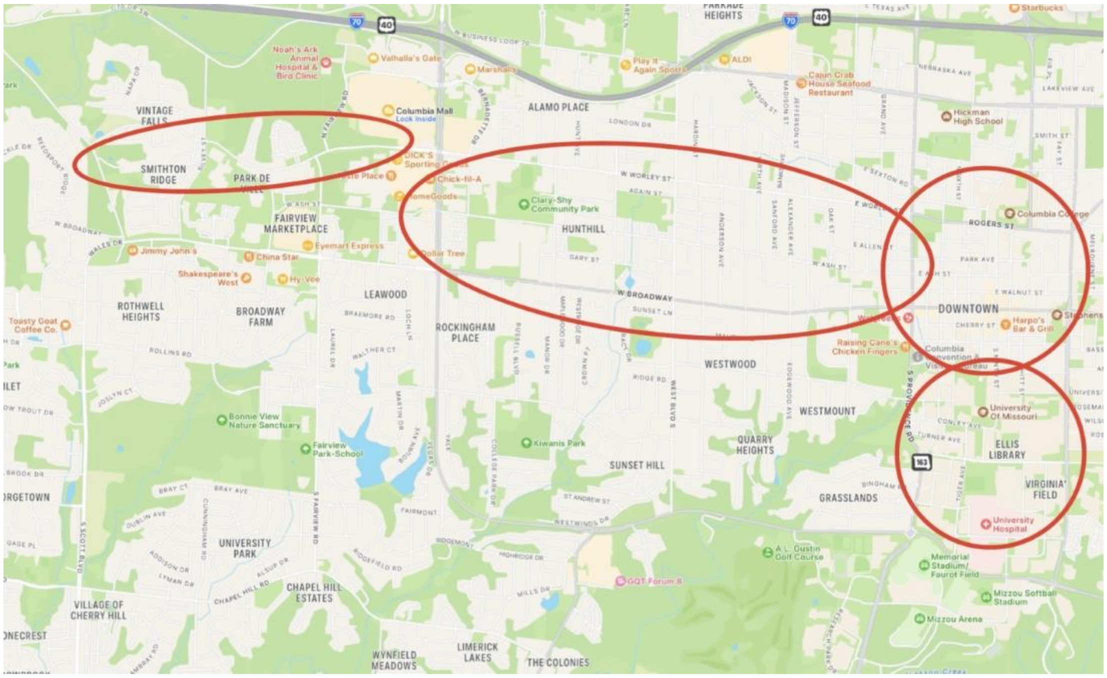


Figure 23. Example Secondary Study Area

Once the data streams were fused, the resulting sign image, sign class, and reflectivity value were then input into a trained large language model (LLM). The LLM generated descriptive phrases for each visual assessment category – legibility, color, shape integrity, and surrounding environment - followed by a calculated overall condition score based on the descriptions and reflectivity, the Sign Condition Index (SCI). The SCI score ranged from 0 to 10, incorporating the visual assessment and the reflectivity performance. For this study, reflectivity carried a higher weight, 70%, of the score than the daytime factors, 30%, emphasizing the importance of the nighttime evaluations. These weights, however, these can be adjusted based on the priorities or specific needs of different municipalities. By automating these assessments, municipalities can quickly locate and prioritize sign replacement or maintenance based on quantifiable conditions ratings.

CHAPTER 7 CONCLUSION

This study successfully developed and demonstrated a novel framework for automated traffic sign condition evaluation that addresses critical challenges faced by transportation agencies. By leveraging vehicle-mounted cameras and LIDAR technology, the research presents a cost-effective, efficient, and scalable approach to monitoring and maintaining traffic sign inventories.

7.1 Research Contributions

The primary contribution of this research is the development of the dual-mode traffic sign evaluation framework that systematically assesses both daytime visual performance and nighttime reflectivity. This approach provides a comprehensive evaluation system that aligns with MUTCD requirements while reducing the resource-intensive nature of traditional inspection methods. Key contributions include:

- **Integration of visual and reflective assessment**: The methodology combines camera-based imagery with LIDAR-derived reflectivity measurements, creating a holistic approach to sign condition evaluation that accounts for both daytime visibility factors and nighttime retroreflectivity performance.
- **Large Language Model application**: The innovative use of a fine-tuned LLM to analyze traffic sign imagery demonstrates the potential of AI-driven solutions in infrastructure management. The model achieved promising performance across evaluation categories.
- **Sign Condition Index (SCI)**: The development of a standardized, quantifiable scoring system (SCI) provides transportation agencies with an objective metric for

prioritizing maintenance activities. This adaptable framework allows for customization based on local priorities and specific operational needs.

- **Automated data processing pipeline**: The research established an end-to-end workflow from data collection to condition assessment, reducing manual effort and increasing the efficiency of sign inventory management.

7.2 Key Findings

Analysis of the 190 traffic signs evaluated in this study revealed several important findings:

- Most traffic signs in the study area were in good condition, with over 50% achieving the maximum SCI score of 10, indicating both excellent visual characteristics and compliant retroreflectivity levels.
- SCI scores ranged from 5.3 to 10, with a mean of 9.31 and a standard deviation of 1.139, suggesting generally well-maintained signage in the Columbia, Missouri area.
- The fine-tuned LLM performed best when evaluating surrounding environment factors (mean $f1$ score of 0.782) and the color characteristics (mean $f1$ score of 0.738), while legibility assessment showed the most room for improvement (mean $f1$ score of 0.660).
- The methodology successfully demonstrated the ability to differentiate between high-quality signs with strong reflectivity and deteriorated signs with faded appearance and reduced reflectivity.

7.3 Practical Implications

The framework developed in this study offers several practical benefits for transportation agencies:

- **Resource optimization**: By automating the sign condition assessment process, agencies can allocate limited resources more efficiently, focusing maintenance efforts on signs that most urgently require attention.
- **Compliance support**: The methodology helps agencies meet MUTCD requirements for sign maintenance through systematic evaluation of both visual characteristics and retroreflectivity levels.
- **Scalable implementation**: The approach is adaptable to various environments and can be implemented by agencies of different sizes, potentially reducing the financial and logistical burden of sign maintenance programs.
- **Enhanced safety**: Improved monitoring of traffic sign conditions contributes to roadway safety by ensuring signs remain visible, legible, and effective at communicating critical information to road users.

7.4 Limitations and Future Research

While the methodology demonstrated strong potential, several limitations were identified that warrant further investigation:

- **Sample distribution**: The limited availability of poor-quality signs in the study area resulted in an imbalanced dataset that may affect model generalization to severely deteriorated signs.

- **Sensor limitations**: The low sampling frequency of the LIDAR sensor occasionally led to minor misalignments between image and point cloud data during vehicle movement.
- **Environmental factors**: Data collection was primarily conducted during favorable weather conditions, leaving questions about model performance during adverse weather or varying lighting conditions.

Future research directions should focus on:

- **Expanding the dataset**: Incorporating a wider range of sign conditions, including severely deteriorated examples, to improve model robustness.
- **Weather resilience**: Testing and adapting the methodology for reliable performance across various weather and lighting conditions.
- **Integration with municipal systems**: Exploring how the automated assessment framework can interface with existing asset management systems to streamline maintenance workflows.
- **SCI threshold calibration**: Developing standardized threshold values based on the Sign Condition Index (SCI) to classify signs into actionable maintenance categories – such as sufficient, needs replacement soon, or requires immediate replacement – tailored to municipal needs and safety standards.

In conclusion, this research demonstrates that combining vehicle-mounted cameras, LIDAR sensors, and machine learning techniques offers a promising approach to automating traffic sign condition assessments. The developed framework addresses key challenges in sign maintenance while providing transportation agencies with practical tools to enhance road safety and optimize resource allocation. As sensing technologies continue to advance and become more accessible, the potential for widespread implementation of automated condition assessment systems will continue to grow, ultimately contributing to safer and more efficiently maintained roadway infrastructure.

REFERENCES

Ai, C., & Tsai, Y. J. (2016). An automated sign retroreflectivity condition evaluation methodology using mobile LIDAR and computer vision. Transportation Research Part C.

Alkhulaifi, A., et al. (2021). Predicting traffic sign retro-reflectivity degradation using deep neural networks. Applied Sciences.

Bai, J., Bai, S., Yang, S., Wang, S., Tan, S., Wang, P., Lin, J., Zhou, C., & Zhou, J. (2023). Qwen-VL: A versatile vision-language model for understanding, localization, text reading, and beyond. arXiv. https://arxiv.org/pdf/2308.12966

City of Los Altos. (n.d.). Traffic sign replacement. https://www.losaltosca.gov/publicworks/project/traffic-sign-replacement

Cloudflare. (2024). What is a large language model (LLM)? https://www.cloudflare.com/learning/ai/what-is-large-language-model/

Federal Highway Administration. (n.d.). Design guidelines. https://mutcd.fhwa.dot.gov/SHSm/design.pdf

Guan, et al. (2020). Integrating mobile LiDAR and digital images for traffic sign inventory and condition assessment.

Idris, I. I., et al. (2024). Machine learning-based framework for prediction of retroreflectivity degradation of pavement markings. Journal of Transportation Engineering.

Jamal, A., et al. (2022). Modeling retroreflectivity degradation of traffic signs using artificial neural networks. IATSS Research.

Kennewick Public Works. (n.d.). Sight obstructions. https://www.go2kennewick.com/368/Sight-Obstructions

Khrapova, M. (2019). Determining the influence of factors on retroreflective properties of traffic signs. Agronomy Research.

Manasreh, D., et al. (2024). Feature-centric approach for learning-based prediction of pavement marking retroreflectivity from mobile LiDAR data. Buildings.

Merolla, D., et al. (2024). Automated road safety: Enhancing sign and surface damage detection with AI.

Rani, A., et al. (2024). Traffic sign detection and recognition using deep learning-based approach with haze removal for autonomous vehicle navigation.

Saleh, R., & Fleyeh, H. (2024). Predictive models for road traffic sign retroreflectivity status. International Journal of Transportation Science and Technology.

Schmitt, J. (2025). Prompt engineering: A guide to improving LLM performance. CircleCI. https://circleci.com/blog/prompt-engineering/

Siegmann, P., et al. (2005). Automatic evaluation of traffic sign visibility using SVM recognition methods.

Steele, A., et al. (2023). Traffic sign retroreflectivity condition assessment and deterioration analysis using lidar technology. Journal of Transportation Engineering.

Swapna, Y., et al. (2021). Deep learning-based road traffic sign detection and recognition.

Transline Inc. (n.d.). What's the lifespan of road signs and what happens then? https://translineinc.com/lifespan-of-road-signs-and-what-happens-then/

You, C., et al. (2017). Rapid traffic sign damage inspection in natural scenes using mobile lidar scanning data.

Zhou, L., & Deng, Z. (2014). LIDAR and vision-based real-time traffic sign detection and recognition algorithm.

Zhu, Z., et al. (2016). Traffic-sign detection and classification in the wild.