

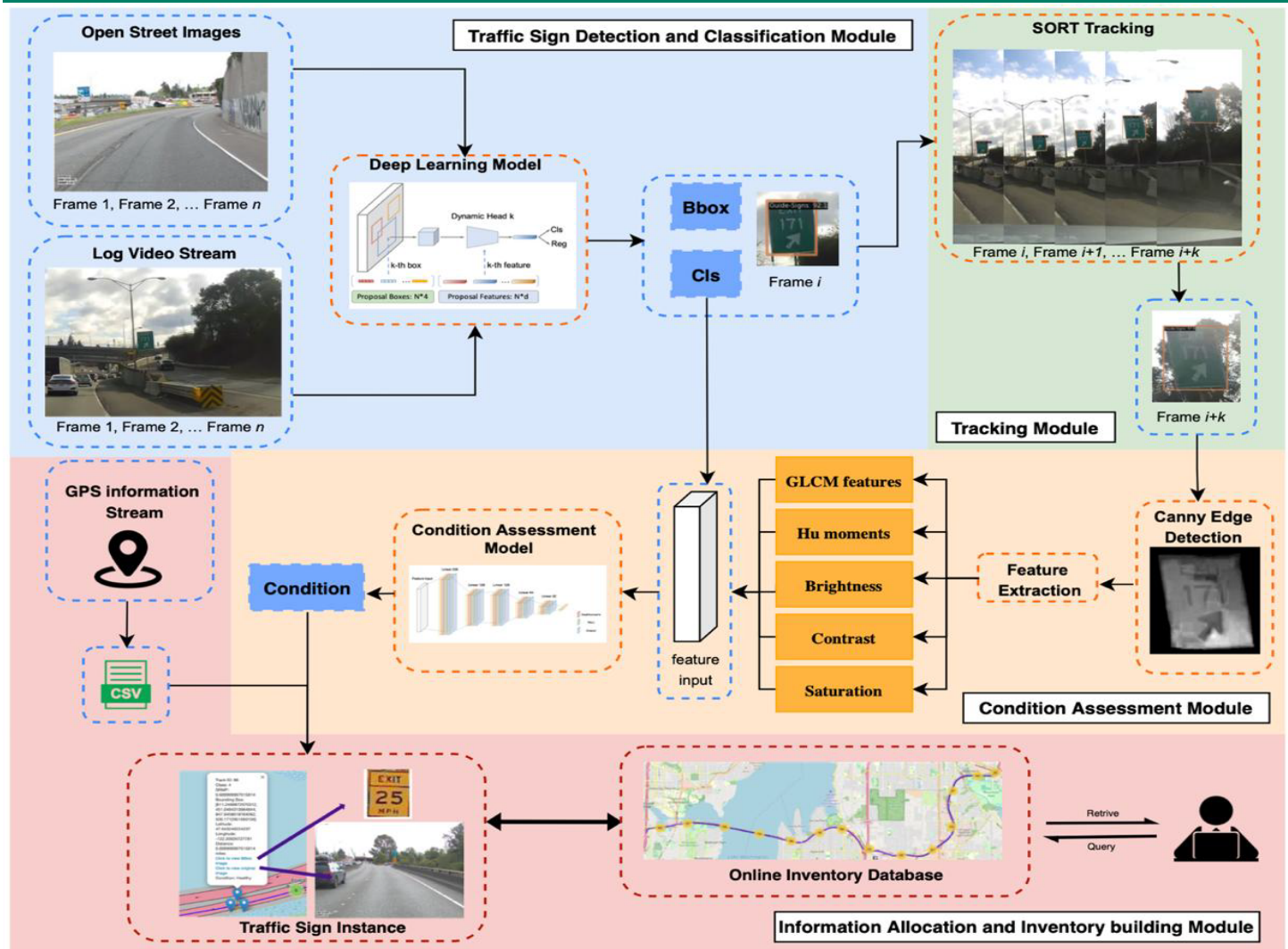
Automated Traffic Sign Recognition Using Computer Vision and Deep Learning

WA-RD 946.1

Yinhai Wang
Chenxi Liu

Shucheng Zhang
Nutvara Jantarathaneewat

March 2025



Washington State
Department of Transportation

Office of Research & Library Services

WSDOT Research Report

Final Research Report
Agreement T1461, Task AA
WSDOT Sign Detection
WA-RD 946.1

Automated Traffic Sign Recognition Using Computer Vision and Deep Learning

by

Yinhai Wang
University of Washington

Shucheng Zhang
University of Washington

Chenxi Liu
University of Washington

Nutvara Jantarathaneewat
University of Washington

Department of Civil and Environmental Engineering
University of Washington, Box 352700
Seattle, Washington 98195

Washington State Transportation Center
(TRAC) University of Washington, Box 359446
University Tower, 4333 Brooklyn Ave NE
Seattle, Washington 98195-9446

Washington State Department of Transportation
Technical Monitors
Dina Swires and Trevor McCain

Prepared for
The State of Washington
Department of Transportation
Julie Meredith, Secretary

March 2025

Technical Report Documentation Page			
1. Report No. WA-RD 946.1		2. Government Accession No.	
4. Title and Subtitle Automated Traffic Sign Recognition Using Computer Vision and Deep Learning		3. Recipient's Catalog No.	
		5. Report Date March 2025	
		6. Performing Organization Code	
7. Author(s) Yinhai Wang, Shucheng Zhang, Chenxi Liu, and Nutvara Jantarathaneewat		8. Performing Organization Report No.	
9. Performing Organization Name and Address Washington State Transportation Center University of Washington, Box 359446 University Tower, 4333 Brooklyn Ave NE Seattle, Washington 98195-9446		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. Contract T1461, Task AA	
12. Sponsoring Organization Name and Address Washington State Department of Transportation Transportation Building, MS 47372 Olympia, Washington 98504-7372 14 Doug Brodin, Project Manager, 360-705-7972		13. Type of Report and Period Covered Final research report	
		14. Sponsoring Agency Code	
15. Supplementary Notes			
16. Abstract An up-to-date traffic sign inventory is essential for roadway safety, infrastructure maintenance, and support of intelligent road agents, yet traditional manual inspections are inefficient and labor-intensive. This research presents an automated system that leverages deep learning and computer vision to detect, classify, and generate a comprehensive traffic sign inventory from log videos and open street images. To enhance model performance, a new traffic sign dataset tailored to Washington state traffic scenarios was developed and used for training. Additionally, traffic sign condition assessment was explored using six retro-reflectivity-related features. The integrated system consists of four modules—detection, tracking, assessment, and inventory building—each designed to perform a specific function, with the final inventory visualized geospatially. For this system, sparse R-CNN was employed for detection and classification, while a DNN classifier was used for assessment. Experimental results demonstrated high detection accuracy (93.50 mAP@50) and reliable condition assessment (93.56 percent accuracy) while maintaining computational efficiency (25.4 FPS). A comparison with the Washington State Department of Transportation (WSDOT) inventory revealed key discrepancies. By automating traffic sign management, this system provides a scalable, cost-effective solution for transportation agencies to maintain accurate inventories, prioritize maintenance, and enhance roadway safety, thereby ensuring more reliable roadway and seamless mobility for all road users			
17. Key Words Traffic sign detection and recognition; traffic asset inventory; deep learning; traffic sign maintenance		18. Distribution Statement No restrictions.	
19. Security Classification (of this report) Unclassified.	20. Security Classification (of this page) Unclassified.	21. No. of Pages 107	22. Price NA

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

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and accuracy of the data presented herein. This document is disseminated through the Washington State Department of Transportation. The contents do not necessarily reflect the views or policies of the Washington State Department of Transportation or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

TABLE OF CONTENTS

ABBREVIATIONS LIST	xi
EXECUTIVE SUMMARY	xiii
Key Products	xiv
1.0 INTRODUCTION AND BACKGROUND	1
2.0 LITERATURE REVIEW	7
2.1 Overview	7
2.2 Traffic Sign Detection and Recognition	8
2.2.1 Hand-Crafted Feature Extraction	8
2.2.2 End-to-End Learning	13
2.3 Traffic Sign Data Collection and Identification Systems	16
2.4 Traffic Sign Condition Assessment	18
3.0 TRAFFIC SIGN DATA SELECTION AND VERIFICATION.....	23
3.1 Overview	23
3.2 Public Traffic Sign Dataset Review.....	23
3.2.1 German Traffic Sign Benchmarks (GTSRB).....	24
3.2.2 Tsinghua-Tencent 100K Dataset.....	24
3.2.3 LISA Traffic Sign Dataset	25
3.2.4 Annotations for Road Traffic Signs (ARTS) Dataset	26
3.3 Open Street Images	26
3.3.1 Google Maps Street View	27
3.3.2 State Route Web Tool	27
3.4 Dataset Selection.....	28
4.0 TRAFFIC SIGN DATA COLLECTION AND DATASET CONSTRUCTION	31
4.1 Overview	31

4.2 Vanilla Traffic Sign Detection Model Training.....	31
4.3 Data Collection System Development	34
4.3.1 System Structure	34
4.3.2 Data Collection Process	36
4.3.3 Post-Processing.....	37
4.4 Dataset Construction.....	39
5.0 LEARNING-BASED TRAFFIC SIGN DETECTION AND RECOGNITION	
ALGORITHM DEVELOPMENT	43
5.1 Overview	43
5.2 Model Selection and Modification.....	43
5.3 Training Process.....	46
5.3.1 Environment Preparation	46
5.3.2 Training Strategies	46
5.4 Performance Comparison.....	49
5.4.1 Evaluation Metrics	49
5.4.2 Results.....	50
6.0 TRAFFIC SIGN CONDITION ASSESSMENT	55
6.1 Overview	55
6.2 Methodologies.....	55
6.2.1 Imagery Assessment Features Selection and Extraction	55
6.2.2 Model Selection	61
6.3 Experiments	61
6.3.1 Methods Comparison	62
6.3.2 Input Features Ablation Study	62
7.0 SYSTEM INTEGRATION AND SAMPLE DATA INVENTORY CREATION	65

7.1 Overview	65
7.2 System Integration	65
7.2.1 Detection and Classification Module.....	66
7.2.2 Tracking Module.....	67
7.2.3 Assessment Module	68
7.2.4 Inventory Building Module.....	68
7.3 Sample Data Inventory Visualization	69
7.4 Comparison of Traffic Sign Inventories for SR 520 and SR 530	76
8.0 CONCLUSIONS AND FUTURE WORK.....	83
8.1 Conclusion	83
8.2 Future Work	85
9.0 REFERENCES.....	87

LIST OF FIGURES

Figure 1 Results of Color-based Traffic Sign Segmentation in the HSI Color Space (Nguwi and Kouzani 2006)	10
Figure 2 Feature Matching Using Three Descriptors	13
Figure 3 Traffic Sign Data Collection System Proposed by Balali et al. (2015).....	17
Figure 4 Sample Traffic Sign Data in GTSRB	24
Figure 5 Sample Traffic Sign Data in the Tsinghua-Tencent 100K Dataset	25
Figure 6 Sample Traffic Sign Data of LISA Dataset.....	25
Figure 7 Sample Traffic Sign Data of ARTS Dataset.....	26
Figure 8 Sample Image from Location (47.636855, -122.323283)	27
Figure 9 SRweb Tool Supported by WSDOT	28
Figure 10 Sample Detection Results by the Faster-R-CNN-Inception-V2-Based TSD Model	34
Figure 11 The Structure of Data Collection System	35
Figure 12 Collection System Overview	36
Figure 13 Data Collection Process.....	37
Figure 14 Labeling Process.....	39
Figure 15 Class Distribution	40
Figure 16 Object Count Per Image	41
Figure 17 Annotation Heatmap.....	42
Figure 18 Sparse RCNN Architecture (Sun et al. 2021).....	45
Figure 19 Sparse R-CNN Loss Changes during Training	49
Figure 20 mAP50 Changes during Training	52
Figure 21 Some Detection Examples.....	54
Figure 22 The Saturation Maps and Histogram of Saturation Value of Stop Signs in Different Ages	58
Figure 23 Brightness and Contrast Heatmaps for Traffic Signs	59
Figure 24 Minimum Maintained Retro-reflectivity Levels in MUTCD.....	60
Figure 25 Assessment Model Architecture.....	61
Figure 26 Examples of Healthy and Damaged Traffic Signs	64
Figure 27 System Architecture	66
Figure 28 ARM and SRMP Locations Along Washington Highways	70
Figure 29 SR-520 and SR-530 Data and Its Locations.....	71
Figure 30 Sample Data Inventory and Cropped Images	71
Figure 31 Inventory Map for SR 520.....	72
Figure 32 Inventory Map for SR 530.....	72
Figure 33 Visualized Inventory for SR 520 (Guide and Information Sign)	73
Figure 34 Visualized Inventory for SR 520 (Warning Sign).....	74
Figure 35 Visualized Inventory for SR 520 (Black and White Regulatory Sign)	74
Figure 36 Visualized Inventory for SR 530 (Route Marker Sign).....	75
Figure 37 Visualized Inventory for SR 530 (Marker)	75

Figure 38 Examples of Detection Results on SR 520.....	76
Figure 39 Comparison of Detected vs Not Detected Traffic Signs	77
Figure 40 A Traffic Sign Instance Both Appearing in WSDOT (Top) and Our Inventories (Bottom)	78
Figure 41 A Digital Speed Limit Sign Only Showing in Our Inventory (Bottom)	79
Figure 42 Traffic Sign Installation Date Comparison.....	79

LIST OF TABLES

Table 1 Public Dataset Comparison.....	24
Table 2 TSD Model Comparison.....	32
Table 3 Vanilla Traffic Sign Detection Model Comparison.....	33
Table 4 Relevant Specifications of the Data Collection Vehicle Dash Camera	34
Table 5 Performance Comparison	51
Table 6 Condition Assessment Models Comparison: Methods.....	62
Table 7 Input Features Ablation Study Results	63

ABBREVIATIONS LIST

ADAS	Advanced driver-assistance system
ARM	Accumulated route mileage
ASCE	American Society of Civil Engineers
BRISK	Binary Robust Invariant Scalable Keypoints
CAV	Connected and autonomous vehicle
CNN	Convolutional neural network
CSV	Comma-separated value
CUDA	Compute Unified Device Architecture
CV	Computer vision
DL	Deep learning
DNN	Deep neural network
DOT	Department of transportation
DT	Distance transform
FPS	Frames per second
GIoU	Generalized Intersection over Union
GIS	Geographic information system
GLCM	Gray-Level Co-Occurrence Matrix
GPS	Global Positioning System
GPU	Graphic processing unit
GSV	Google Maps Street View
GTSRB	German Traffic Sign Recognition Benchmark
HSI	Hue-saturation-intensity
HSV	Hue-saturation-value
IoU	Intersection over Union
ITS	Intelligent transportation systems
LiDAR	Light detection and ranging
MAE	Mean absolute error
mAP	Mean average precision
MUTCD	Manual on Uniform Traffic Control Devices
RAM	Random-access memory

RGB	Red-green-blue
ROI	Region of interest
SRMP	State route milepost
SURF	Speeded Up Robust Features
STAR Lab	Smart Transportation Applications & Research Laboratory
TSD	Traffic sign detection
WSDOT	Washington State Department of Transportation

EXECUTIVE SUMMARY

This study presents an automated approach to traffic sign detection, recognition, and condition assessment using computer vision and deep learning methodologies. The proposed system streamlines traffic sign inventory management, significantly improving efficiency over traditional manual and semi-automated methods. By integrating public datasets, self-collected data, and open-source imagery, the system will ensure adaptability to various environmental conditions, enhancing real-world applicability.

A key component of the system is the deployment of a learning-based traffic sign detection and recognition algorithm, for which Sparse R-CNN was selected because of its superior balance between accuracy and computational efficiency. The model successfully identifies traffic signs in challenging environments, demonstrating its robustness for real-world deployment. Additionally, an advanced image-based condition assessment method was developed to evaluate maintenance status. The developed system was integrated into a traffic sign inventory management framework, incorporating GPS data, classification results, and condition assessments. This interactive inventory enables real-time monitoring, query-based retrieval, and historical tracking of traffic signs, which will allow transportation agencies to streamline infrastructure management and optimize resource allocation.

While the research delivers a scalable and cost-effective solution for automated traffic sign management, future improvements will focus on enhancing small object detection, optimizing real-time deployment, incorporating additional sensor data for condition assessment, and integrating with GIS for enhanced spatial analysis.

The findings and technological innovations presented in this study will contribute to enhancing road safety, regulatory compliance, and efficient transportation asset management in Washington state and beyond.

KEY PRODUCTS

The research produced several tangible outputs, including the following:

- A data collection system for traffic sign imagery and related information.
- A software solution for traffic sign detection, recognition, and classification.
- A mapping system for visualizing traffic signs on road networks.
- A user guide for the developed system.
- A sample traffic sign data inventory.
- A system demonstration showcasing the data inventory creation process.

1.0 INTRODUCTION AND BACKGROUND

In 2017, the U.S. invested over \$177 billion in roads and bridges to enhance their condition and safety (ASCE, 2021). Despite this substantial investment, a \$786 billion backlog remains, primarily due to increasing maintenance demands. This challenge is expected to worsen as infrastructure continues to age and deteriorate. Among various road assets, traffic signs require more frequent monitoring and maintenance because of their relatively short lifespan in comparison to other infrastructure components (Temel et al., 2019). They play a vital role in providing critical information to all road users, including restrictions, warnings, navigation, and updates on road conditions, making them indispensable for smooth traffic operations. Ensuring their accuracy and visibility is essential for efficient traffic flow. Beyond aiding human drivers, traffic signs are increasingly crucial for the operation of emerging technologies such as connected and autonomous vehicles (CAVs). These systems depend on clear, consistent, and up-to-date signage to navigate complex environments safely (Alnasser et al., 2019).

However, traffic signs often suffer from the lack of effective management, leading to their rapid deterioration from factors such as vandalism, insufficient maintenance, and natural degradation (Ai et al., 2016). Research by Boggs et al. (2013) showed that damage to signs significantly reduces their clarity, affecting their visibility both day and night. Specifically, vandalism has been identified as a major cause of sign damage and the only one caused by humans. Khalilikhah et al.'s study (2016) highlighted this issue, showing that around 7 percent of 97,314 surveyed traffic signs in Utah were damaged, with 27 percent affected by vandalism. The situation is getting worse. Seattle Department of Transportation spends over \$1.4 million to clean up graffiti along roads in the greater Seattle area every year (KOMO News, 2022) because it could cover the necessary information or even cause misguidance.

In addition to human activities, natural wear and tear is another significant factor in deterioration (You et al. 2017). Environmental factors such as weather conditions, exposure to sunlight, and seasonal changes can gradually degrade the materials of traffic signs, diminishing their visibility over time. For example, the Louisiana Department of Transportation found that over half of the signs beyond their 10-year warranty failed to meet minimum visibility standards and needed to be replaced immediately (Ré et al. 2011). Similarly, the Connecticut Department of Transportation in its assessment report (CTDOT 2022) highlighted that 59.4 percent of 263,000 signs were in poor condition and marked this situation as high risk.

Arizona Department of Transportation, in its Roadway Report (ADOT 2019), assumed that the maintenance costs for sign panels per mile was \$15k. Given the complexity of the roadway network, the annual costs of maintaining and replacing damaged traffic signs are expected to be huge. However, if damaged traffic signs are not repaired promptly, they can compromise road safety for all users by providing unclear or invisible information, increasing the risk of serious traffic incidents (Magnussen et al., 2020). Missing or illegible signs can lead to driver confusion, misinterpretation of road conditions, and delayed decision-making, which may result in accidents, especially in high-risk areas such as at intersections, on highways, and in construction zones. Ensuring timely maintenance of traffic signs is, therefore, critical to maintaining safe and efficient transportation systems.

To efficiently allocate limited resources for timely maintenance, an up-to-date and comprehensive traffic sign inventory is essential. This inventory should systematically record detailed information about each traffic sign, including its location, type, installation date, material composition, condition, and visibility status. With such a well-maintained inventory, traffic asset management can be streamlined, enabling transportation agencies to prioritize

repairs and replacements based on urgency and available resources. Additionally, a robust inventory facilitates high-level maintenance planning by identifying trends in sign deterioration, optimizing inspection schedules, and ensuring compliance with safety regulations. Ultimately, an effective traffic sign management system enhances transportation efficiency and safety by reducing the risk of outdated, damaged, or missing signage that could compromise road users' navigation and decision-making. However, the current traffic sign inventories maintained by transportation agencies often suffer from outdated records, missing signs, and inconsistent updates. Without a reliable and frequently updated inventory, road safety can be compromised, and the effectiveness of advanced driver-assistance systems (ADASs) and CAV technology may be significantly reduced. Therefore, developing an automated and scalable approach for traffic sign inventory generation is essential to managing transportation infrastructure and ensuring that road assets remain in optimal condition.

Traditional methods for traffic sign inventory collection involve manual assessment, which is labor-intensive, time-consuming, and prone to human error. While some agencies have adopted semi-automated methods using Global Positioning Systems (GPS) and image-based logging, these approaches still require significant human oversight and are not scalable for large-scale deployments (Radoš et al. 2022). Recent advancements in computer vision (CV) and deep learning (DL) have demonstrated strong potential for automating traffic sign detection, classification, and condition assessment. By leveraging these technologies, traffic sign inventory management can become more accurate, efficient, and cost effective. However, two key challenges must be addressed to develop and implement such systems successfully. The first challenge is obtaining reliable ground-truth data, including the precise location, type, and

condition of each traffic sign. The second involves deploying algorithms and systems capable of performing this process with high accuracy and efficiency.

Therefore, the primary goals of this project were to establish an automated pipeline for traffic sign data collection and identification using vision-based deep learning methods and to develop a sample traffic sign data inventory for Washington state. The tasks and achievements in this project are summarized in the following four key points:

- **Develop a deep-learning-based approach for automated traffic sign detection and recognition:** This report presents a robust and accurate deep-learning framework for traffic sign detection and recognition from open street images. Unlike traditional CV methods that rely on hand-crafted feature extraction, our approach leverages advanced deep-learning techniques to handle diverse weather conditions and complex environments.
- **Create an end-to-end pipeline from traffic sign data collection to inventory generation:** We designed and implemented a fully integrated pipeline that automates the process of collecting and identifying traffic signs from open street images or log videos and then organizes them into a dynamic inventory. This pipeline addresses a critical gap in current methodologies by seamlessly combining data collection and recognition tasks, eliminating the need for manual intervention and significantly improving efficiency.
- **Establish a sample traffic sign data inventory in Washington state:** We generated a sample inventory of traffic signs extracted from publicly available imagery, complete with associated metadata. This inventory is designed to support WSDOT's transportation asset management plans and can be periodically updated and expanded

by using the same automated workflow, ensuring scalability and long-term applicability.

- **Generate a new U.S.- based traffic sign dataset:** A significant challenge hindering the advancement of intelligent transportation systems (ITS) in the U.S. is the lack of a comprehensive traffic sign dataset. To address this gap, we developed a new dataset primarily composed of traffic sign images collected from various locations across Washington state. This dataset captures a diverse range of traffic signs under varying weather conditions, lighting scenarios, and environmental factors, enhancing the robustness and generalizability of traffic sign recognition models. Our recognition model, trained on this dataset, demonstrated strong performance and adaptability in field tests.

By utilizing advanced object detection and tracking models, feature-based condition assessment techniques, and interactive inventory visualization tools, this research provides a scalable and efficient solution for building and maintaining an up-to-date traffic sign inventory. The proposed system can assist transportation agencies in proactively identifying damaged or missing signs, prioritizing maintenance efforts, and enhancing roadway safety for both human drivers and autonomous systems.

2.0 LITERATURE REVIEW

2.1 OVERVIEW

This literature review highlights the progression of traffic sign detection and recognition technologies, from traditional, hand-crafted feature extraction methods to modern, end-to-end deep learning models. Early approaches relied on color- and shape-based detection techniques, laying the groundwork for traffic sign recognition. However, their limitations in handling complex, real-world scenarios, such as varying lighting conditions, occlusions, and background clutter, prompted a shift toward deep learning-based solutions. Convolutional neural networks (CNNs) and other architectures have achieved state-of-the-art performance. These advancements have made real-time, high-accuracy traffic sign recognition possible, especially for autonomous driving applications.

Beyond detection and recognition, integrating these models into traffic sign data collection and identification systems has become crucial for practical deployment. Automated traffic sign inventory management has demonstrated feasibility, offering more efficient monitoring and maintenance solutions. However, challenges remain in accommodating regional variations in sign designs, materials, and environmental conditions, as well as in scaling these systems to meet the operational needs of transportation agencies.

Additionally, the review highlights the increasing importance of traffic sign condition assessment. Traditional manual inspection methods are labor-intensive, inefficient, and costly, driving the demand for automated assessment systems. Emerging solutions, including light detection and ranging (LiDAR), vision-based methods, and deep learning, have shown promise for detecting damaged, missing, or deteriorated signs. However, challenges such as limited

annotated datasets, high implementation costs, and seamless system integration must be addressed to enhance reliability and scalability.

2.2 TRAFFIC SIGN DETECTION AND RECOGNITION

Traffic sign classification and detection have evolved through two main phases: hand-crafted feature extraction and end-to-end learning. The hand-crafted feature extraction phase primarily relied on manual selection and engineering of features such as color, shape, texture, and geometric properties. These methods often used traditional image processing and pattern recognition techniques to identify and classify traffic signs. Although effective in controlled conditions, these approaches struggled with variations in illumination, occlusions, and complex backgrounds, necessitating extensive preprocessing and parameter tuning to achieve acceptable performance. In contrast, the advent of end-to-end learning brought a paradigm shift, with deep learning models automatically extracting relevant features from raw image data. By leveraging large datasets and powerful computational resources, these models have demonstrated remarkable accuracy and robustness, even in challenging scenarios. End-to-end learning techniques, particularly those based on CNNs, have become increasingly popular for traffic sign detection and classification due to their scalability and adaptability across diverse conditions.

2.2.1 Hand-Crafted Feature Extraction

Color-Based Detection

Color-based detection techniques rely on color as the primary feature for identifying regions in images or video frames that contain traffic signs. A common method within this category is color segmentation, which removes unnecessary objects, thereby reducing the search area in the image or frame. Supreeth et al. (2016) proposed a method for detecting red traffic signs in video frames using color as the main feature. In this approach, the original video frame

is first converted to grayscale, and only the red channel is extracted. The grayscale frame is subtracted from the red channel of the original frame, and the resulting difference frame is converted to binary using a threshold. Qin et al. (2010) discussed the color distance method for detecting traffic signs from image sequences. This method calculates the color distance, like the Euclidean distance between two points, by measuring the difference between two colors. As the color distance decreases, the similarity between colors increases. A color standardization method was proposed by Xing et al. (2016) to detect traffic signs. In this approach, each pixel in the red-green-blue (RGB) space is compared against threshold values defined for each component. After standardization, the image contains only black, red, green, blue, cyan, magenta, and white. The process retains only the relevant colors for traffic signs, converting the image into a binary format.

The above methods utilize the RGB color space for segmentation. While RGB segmentation is computationally efficient, it is sensitive to changes in illumination. To address this limitation, Nguwi and Kouzani (2006) utilized the Hue-Saturation-Intensity (HSI) color space for traffic sign detection. By separating achromatic and chromatic components, the HSI color space enables detection under challenging lighting conditions. An example of stop sign detection using this method is shown in Figure 1.

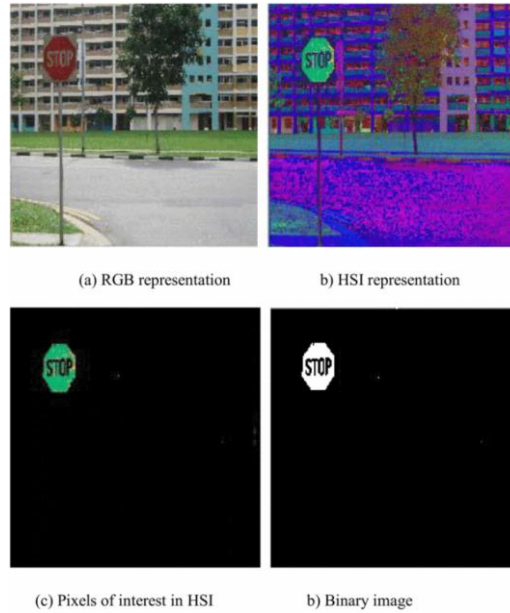


Figure 1 Results of Color-based Traffic Sign Segmentation in the HSI Color Space (Nguwi and Kouzani 2006)

Similarly, Romdhane et al. (2016) used the Hue-Saturation-Value (HSV) color space for traffic sign detection. After the RGB image is converted to HSV, segmentation is conducted on the basis of the Hue and Saturation values. Background noise is removed through morphological operations, such as erosion, dilation, opening, and closing. Connected component analysis is then applied to identify potential traffic sign regions that meet predefined constraints for height, width, and area. Fatmehsari et al. (2010) proposed that the YCbCr color space be employed for traffic sign detection. The red component is extracted by applying dynamic thresholding to the Cr component. Objects with areas smaller than a predefined threshold are discarded, as they are unlikely to be traffic signs.

Shape-Based Detection

A significant challenge in color-based detection methods is their sensitivity to illumination. To address this limitation, shape-based detection methods were introduced. Traffic signs typically have distinct shapes, such as triangles, rectangles, octagons, and circles, making

shape analysis a practical approach for detection. This section discusses some widely used algorithms for shape-based detection.

A template-based method was proposed by Garvila (1999) for traffic sign detection. This approach involves two binary images: a feature image and a feature template. The distance transform (DT) of the feature image is used to measure the matching between T and I. In the DT image, each pixel value represents the distance of that pixel to its nearest edge. The template is matched against the DT image to identify the shape of interest. While effective, this method comes with a high computational cost because of the template-matching process. Gracia-Garrido et al. (2006) used the Hough transform to detect traffic signs. Circular traffic signs are identified by using the Hough transform for circumferences, while triangular signs are detected by using the Hough transform for straight lines. The process begins with edge detection using the Canny edge detection method. However, analyzing all the contours in an image can result in high computational costs. To optimize performance, contours are filtered on the basis of their area and perimeter, reducing the number of computations. The Hough transform is then applied to the filtered contours, achieving detection rates of 97.2 percent for speed limit signs and 94.3 percent for warning signs.

Similarly, Barnes and Zelinsky (2004) used the radial symmetry transform. The original image is converted to grayscale, and the gradient of each pixel is computed. Two “vote images,” containing gradient orientation and magnitude information, are constructed and referred to as the orientation and magnitude images. A radial symmetry image is then formed by combining these two vote images. Circles are detected by applying a threshold to the radial symmetry image. Unfortunately, the radial symmetry algorithm is limited to detecting regular polygons and is not effective for traffic signs with geometric distortions.

Feature Matching Recognition

After traffic signs are detected in images or video frames, the next step is to recognize which specific traffic sign it is. Recognition can be performed by using feature-matching algorithms. Zheng et al. (2012) utilized Scale and Rotation Invariant Binary Robust Invariant Scalable Keypoints (BRISK) for traffic sign recognition. BRISK is faster than the widely used Speeded Up Robust Features (SURF) algorithm, making it a more efficient choice for real-time applications. BRISK operates in two main stages: keypoint detection and binary descriptor extraction. In the first stage, points of interest are identified in the image. In the second stage, pixel-wise comparisons of brightness are used to generate a binary bit-string, forming the descriptor vector. The template image with the highest number of matching keypoints is classified as the target traffic sign.

Oruklu et al. (2012) employed an interest point detector based on the SURF algorithm for traffic sign recognition. SURF serves as both a detector and a descriptor. As a detector, it identifies interest points, such as blobs and corners, within the image. As a descriptor, it forms a feature vector representing these detected interest points. The feature vectors of the target traffic sign are then compared to those stored in a database to identify the best similarity match. Figure 2 shows feature matching using three descriptors based on above algorithms.

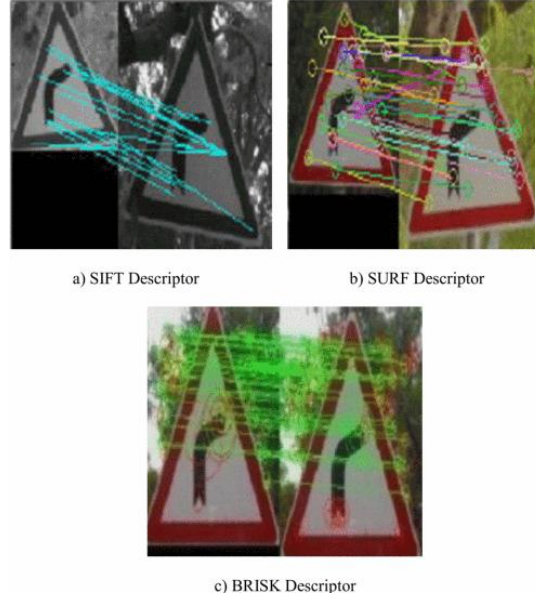


Figure 2 Feature Matching Using Three Descriptors

2.2.2 End-to-End Learning

Recently, deep learning-based object detection methods have gained wide attention as a result of their ability to learn deep features for more accurate localization and classification of objects. Unlike traditional methods that rely on hand-crafted features, deep learning models leverage neural networks to learn directly from raw data, leading to better accuracy and robustness across diverse scenarios. Recent advancements have demonstrated that end-to-end learning approaches not only simplify the workflow but also enhance performance under challenging conditions such as varying illumination, partial occlusions, and background complexity. This section explores notable contributions to this field, focusing on ten significant studies that highlight the power of deep learning in traffic sign detection recognition.

In 2017, a pivotal contribution by Shustanov and Yakimov (2017) introduced a CNN architecture specifically designed for traffic sign recognition. Their model, comprising three convolutional layers and a softmax classifier, was evaluated on the German Traffic Sign Recognition Benchmark (GTSRB) dataset, achieving a near-perfect accuracy of 99.94 percent.

This study demonstrated that increasing network depth and optimizing hyperparameters could significantly improve recognition accuracy, even on complex datasets with varying traffic sign appearances. Then, the efficiency of CNNs was further explored by Bangquan and Xiong (2019) through the development of ENet, a hybrid architecture combining elements of VGG16 and LeNet. Tested on the GTSRB dataset, ENet achieved accuracy rates of 98.6 percent with LeNet and 96.7 percent with VGG16. While LeNet's implementation was slower and more resource-intensive, it provided higher accuracy, making it suitable for applications where precision is paramount. This study highlighted the trade-offs between computational efficiency and model accuracy in traffic sign recognition.

Zaibi et al. (2021) enhanced the LeNet-5 architecture by incorporating batch normalization and dropout layers to improve training speed and stability. Tested on the GTSRB dataset, their model achieved an accuracy of 99.84 percent. This research showcased how minor architectural modifications, such as normalization and regularization techniques, can significantly improve model performance, particularly with large datasets containing diverse traffic sign classes. Zhu and Yan (2022) explored two advanced object detection frameworks, YOLOv5 and SSD, for traffic sign recognition. YOLOv5 achieved an accuracy of 97.70 percent, outperforming SSD's 90.14 percent on a self-collected dataset. YOLOv5's ability to process images in real time while maintaining high accuracy makes it ideal for autonomous vehicle applications. The study also emphasized the importance of data augmentation for improving the generalization of deep learning models, showcasing how real-time detection frameworks are essential for practical deployment.

To achieve higher accuracy, the state-of-the-art models are inevitably getting more and more complex. However, computational resources are limited, especially for mobile devices.

Building on the need for lightweight architectures, Li and Wang (2018) utilized MobileNet, a convolutional neural network optimized for mobile and embedded systems. Their approach incorporated batch normalization and ReLU activation to enhance training stability and performance. Tested on the GTSRB dataset, the model achieved an impressive accuracy of 99.66 percent. To further improve robustness, data augmentation techniques such as rotation, scaling, and shifting were applied, demonstrating the effectiveness of MobileNet for real-time, resource-constrained environments. Lightweight models for traffic sign recognition were further refined by Zhang et al. (2020), who introduced a teacher-student network. This approach utilized a teacher network with deep layers to train a smaller, student network designed for deployment in resource-constrained environments. By applying pruning techniques to reduce computational costs, the student network achieved 99.38 percent accuracy on the GTSRB dataset, demonstrating the potential of model compression for real-world applications.

These advancements in end-to-end learning illustrate the transformative impact of deep learning on traffic sign recognition. Models have progressed from simple CNN architectures to highly optimized systems capable of real-time detection and classification across diverse datasets. However, over the past two decades, learning-based traffic sign detection and recognition models have primarily been designed to support advanced driver-assistance systems (ADAS) and CAVs, with less emphasis on applications that aid transportation agencies and governments in managing infrastructure. This project aimed to bridge that gap by developing models tailored for traffic sign management and maintenance, enabling governments to manage these critical assets more efficiently and effectively.

2.3 TRAFFIC SIGN DATA COLLECTION AND IDENTIFICATION SYSTEMS

While advancements in traffic sign detection and recognition models have significantly improved accuracy and efficiency, the practical deployment of these models requires robust systems for data collection, storage, and management. A comprehensive traffic sign data collection and identification system is essential for integrating these models into real-world transportation infrastructure. Such systems enable the continuous monitoring and updating of traffic sign inventories, providing agencies with actionable insights for effective management and maintenance. Recent studies have explored various approaches to designing these systems, focusing on scalable data collection methods, region-specific customization, and leveraging state-of-the-art deep learning techniques.

Automated data collection systems are important for accommodating advanced recognition models. Balali et al. (2015) proposed a system that utilizes Google Street View images to extract and manage traffic sign data (Figure 3). By applying computer vision techniques, this approach automates the detection, classification, and geolocation of traffic signs. The system provides a scalable and cost-effective solution for maintaining up-to-date inventories without the need for extensive field surveys, demonstrating the potential for large-scale traffic sign management systems.

Traffic sign designs and regulations vary significantly across regions, necessitating customized systems for accurate detection and identification. Andreas et al. (2015) emphasized the importance of tailoring traffic sign recognition algorithms to accommodate the diversity in U.S. traffic signs. Their work highlighted the need for systems that can adapt to regional variations, ensuring accurate detection and classification while adhering to local regulations.

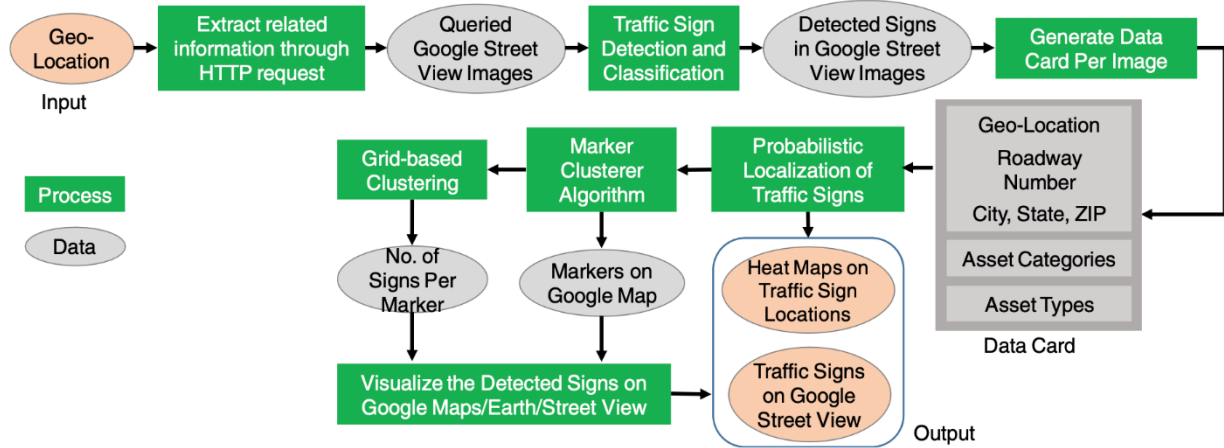


Figure 3 Traffic Sign Data Collection System Proposed by Balali et al. (2015)

Deep learning has transformed traffic sign identification by enabling robust and accurate recognition. However, deploying these models requires systems capable of integrating their computational requirements with real-world applications. Fernando and Sotheeswaran (2021) utilized the YOLOv4 network to create a real-time traffic sign identification system. Their approach achieved an average accuracy of 84.7 percent across 43 traffic sign classes, showcasing the suitability of YOLOv4 for fast and efficient recognition tasks. Such systems highlight the importance of balancing speed and accuracy for real-time applications in traffic management.

Despite such progress, the integration of traffic sign detection and recognition models into real-world systems presents several challenges. Variations in sign appearance, environmental conditions, and dataset generalizability are persistent issues. Additionally, developing lightweight, scalable systems that accommodate real-time processing requirements while maintaining high accuracy remains a critical need. Future efforts should focus on designing systems with robust data pipelines, enabling seamless integration of advanced models for traffic sign detection, identification, and management.

2.4 TRAFFIC SIGN CONDITION ASSESSMENT

With the development of advanced sensors, many researchers started to deploy automated sensing systems to assess traffic signs. Previous research on damaged traffic sign detection and assessment primarily utilized either 3D point cloud analysis to measure retro-reflectivity. Initially, point cloud data were used for detecting physical deformations such as bending and cracking because of their spatial information. However, Ai and Tsai (2016) introduced a new lidar-based approach, demonstrating the feasibility of establishing a relationship between retro-intensity values and retro-reflectivity conditions, regardless of color or shape. González et al. (2011) developed a numerical correlation in their subsequent study, enabling the determination of minimum retro-intensity thresholds for the condition assessment of different signs, turning this method into practice. Similarly, conventional vision-based methods, which had been developed earlier, were designed to match correlations between image intensity and retro-reflectivity. Siegmann et al. (2008) systematically proved that first in their research. Using established correlations, several image-based traffic sign retro-reflectivity condition assessment systems have emerged, including SMARTS, RetroView, VISUALISE, and AMAC (Ai and Tsai 2016). Balali et al. (2019) presented an algorithm that can measure retro-reflectivity during the daytime by simulating the night timing visibility of traffic signs. However, despite their more reliable and consistent condition assessment results in comparison to current retro-reflectometer measurements, the inherent problems of reliance on retro-reflectivity detection are still unavoidable.

- **The correlation between retro-intensity values and retro-reflectivity should be labeled individually in terms of different traffic signs.** The variety in color and shape of traffic signs leads to different patterns of retro-reflectivity. This variability

offers a potential method for classifying traffic signs. However, there is no comprehensive model that could handle all kinds of traffic signs, which means that each category must be individually labeled and measured. Additionally, it's important to note that retro-reflectivity changes with the distance and angle from which it is measured. Therefore, to ensure accuracy, we must establish another correlation between the retro-reflectivity of the same sign observed at different distances and angles before this method can be practically applied. Therefore, the early-stage work will be substantial and time-consuming.

- **Most of these methods have been tested on a limited number of samples within specific traffic sign categories.** For instance, Ai's method was applied only to stop signs, with testing on just 35 samples (Ai and Tsai 2016). Despite these samples representing different conditions, the limited scale makes it difficult to believe that the method is robust and can be generalized to other categories. Furthermore, the fact that none of these methods has been widely adopted by a DOT may prove their limitations.
- **The operation cost is relatively high, and the application environments are limited.** Despite the effectiveness of these point cloud methods, they come with the drawback of requiring expensive laser scanning hardware and computational resources. Similarly, although the vision-based methods are much cheaper, most of them are limited to detecting retro-intensity exclusively at night using an active infrared illuminator, which increases their labor costs.

By applying the robust analytical capabilities of deep learning to imagery data, researchers have demonstrated the significant potential of assessing the condition of traffic signs

from an imagery-based perspective. Deep learning methods offer the ability to analyze large-scale datasets, making them well-suited for traffic sign condition assessment. Additionally, these methods can work under diverse lighting and weather conditions, overcoming the limitations of traditional retro-intensity detection, which often requires specific environments, such as nighttime or active infrared illumination. Therefore, by integrating deep learning into traffic sign assessment, researchers can develop scalable and efficient solutions that eliminate the need for individual retro-reflectivity labeling or distance-specific calibrations. This not only streamlines the early-stage work but also enhances the generalizability of condition assessment across various traffic sign types and categories. Acilo et al. (2018) utilized a ResNet-50 classifier for dual binary classification tasks with images: assessing compliance with design standards and determining the condition of signs as “good” or “bad.” Similarly, Chen and Ren (2023) developed a MFL-YOLO network to identify traffic sign damage, achieving 93.3 percent mAP on their dataset. Rados et al. (2022) extended the EfficientDet framework to detect occlusion-related damages that affect sign readability.

However, these methods still have limitations.

- **Lack of well-established datasets for damage annotation:** One of the primary challenges in developing a robust traffic sign assessment model is the absence of comprehensive datasets that annotate damage conditions. Widely used datasets, such as GTSDb, TT-100K, LISA, and MTSD (Ertler et al. 2022), encompass diverse, country-specific traffic sign scenarios but lack detailed labels regarding sign conditions. This limitation makes them insufficient for training models specifically designed for traffic sign condition assessment.

- **Dependence on costly 3-D data for sign evaluation:** Many existing assessment systems rely on 3D data points, such as LiDAR, to evaluate traffic sign conditions. While effective, these methods are often expensive, computationally intensive, and slow to implement at scale. As a cost-effective and efficient alternative, vision-based approaches utilizing standard imagery provide a promising solution for automated sign assessment.

To address these gaps, this project also focused on developing a new dataset that explicitly labels the conditions of traffic signs, enabling more effective training of assessment models. Additionally, we designed a feature extraction pipeline that leverages image-based techniques to evaluate traffic sign conditions from a visual perspective. By implementing this approach, we aimed to enhance the accuracy, efficiency, and scalability of automated traffic sign condition monitoring systems.

3.0 TRAFFIC SIGN DATA SELECTION AND VERIFICATION

3.1 OVERVIEW

This section outlines the process of identifying and comparing datasets to support the development of a robust traffic sign detection and recognition model. To ensure the effectiveness of the models, this research leveraged a combination of public datasets, open street imagery, and self-collected data. This multi-source approach aimed to address the limitations of existing datasets while ensuring comprehensive coverage of traffic sign types and road conditions.

The section begins with a review of four high-quality public datasets—GTSRB, Tsinghua-Tencent 100K, LISA Traffic Sign Dataset, and ARTS Dataset. It then explores open street imagery as a valuable resource for traffic sign data collection, focusing on the capabilities of tools such as Google Maps Street View (GSV) and the State Route Web Tool (SRweb) for extracting and analyzing real-world traffic sign information. Lastly, the section details the selection process and rationale for organizing datasets into two groups to support the training of a vanilla detection model for data collection and a high-precision detection and recognition model for deployment.

3.2 PUBLIC TRAFFIC SIGN DATASET REVIEW

Traffic sign data collection and identification have been studied extensively, leading to the creation of several public datasets. However, many available datasets lacked the quality or characteristics necessary for this project. Therefore, we selected some instances from four high-quality datasets for model training and validation: the German Traffic Sign Recognition Benchmarks (GTSRB) (Houben et al. 2013), Tsinghua-Tencent 100K (Zhu et al. 2016), the LISA Traffic Sign Dataset (Philipsen et al. 2015), and the Annotations for Road Traffic Signs (ARTS) Dataset (Almutairy et al. 2019). Table 1 shows a comparison of the four public datasets.

Table 1 Public Dataset Comparison

Dataset	Resolution	Number of Images	Number of Classes	Usage	Country
GTSDB	1360 × 800	900	3	Detection	Germany
GTSRB	15 × 15 to 250 × 250	Over 50,000	43	Classification	Germany
ARTS	1920 × 1080	16023	175	Detection	USA
LISA	640 × 640 to 1024 × 522	6,610	47	Detection	USA
TT100K	2048 × 2048	100,000	221	Detection	China

3.2.1 German Traffic Sign Benchmarks (GTSRB)

The GTSRB dataset is a widely used benchmark for traffic sign recognition. It contains over 50,000 images of traffic signs captured in Germany, labeled into 43 classes, including common signs such as stop and speed limit signs, as well as specialized ones such as Roundabout Mandatory (Figure 4). The dataset is divided into a training set (39,209 images) and a testing set (12,630 images), with each image resized to 32x32 pixels. One of its unique strengths is the diversity of lighting, weather conditions, angles, and distances, making it suitable for testing algorithms in real-world scenarios. Because of its high quality and variability, GTSRB has been extensively used in evaluating deep learning models such as CNNs for traffic sign detection and classification.

**Figure 4 Sample Traffic Sign Data in GTSRB**

3.2.2 Tsinghua-Tencent 100K Dataset

The Tsinghua-Tencent 100K dataset contains over 100,000 images with fine-grained annotations (Figure 5). It was created through a collaboration between Tsinghua University and Tencent AI Lab. Designed as a benchmark for image recognition algorithms, it includes rich attribute annotations such as color, shape, texture, and material. The dataset also incorporates

visually similar images with subtle differences, challenging models to distinguish fine-grained attributes effectively. Its large size and detailed annotations provide significant value for training robust recognition systems.

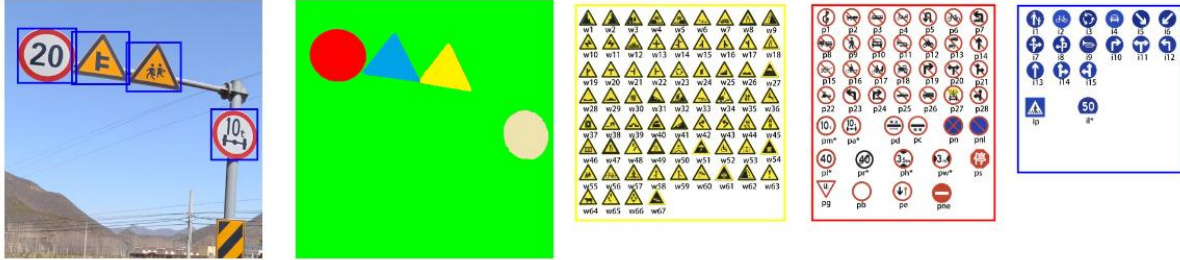


Figure 5 Sample Traffic Sign Data in the Tsinghua-Tencent 100K Dataset

3.2.3 LISA Traffic Sign Dataset

The LISA Traffic Sign Dataset, created by the Laboratory for Intelligent and Safe Automobiles, contains over 6,000 images of U.S. traffic signs divided into 47 classes. It includes a diverse set of lighting and weather conditions, as well as variations in angles and distances (Figure 6). These features make it particularly relevant for U.S. transportation scenarios, enabling models trained on it to perform effectively in real-world environments. The dataset also includes detailed annotations specifying the location, size, and type of each sign, offering valuable information for training and testing algorithms.



Figure 6 Sample Traffic Sign Data of LISA Dataset

3.2.4 Annotations for Road Traffic Signs (ARTS) Dataset

The ARTS Dataset is a large-scale dataset specifically designed for autonomous driving and traffic sign research. It contains over 1.1 million annotated traffic signs from various road environments, including urban, suburban, and rural areas. One of its key features is the inclusion of signs captured under diverse environmental conditions, such as varying weather, illumination, and occlusion levels (Figure 7). ARTS provides comprehensive annotations, including bounding boxes, classes, and hierarchical label structures, enabling multi-task learning approaches. Its scale and diversity make it an invaluable resource for training deep learning models that need to generalize well across different regions and scenarios.



Figure 7 Sample Traffic Sign Data of ARTS Dataset

3.3 OPEN STREET IMAGES

In addition to public traffic sign datasets, open street imagery offers a valuable resource for traffic sign data collection and identification. Tools such as GSV and SRweb provide extensive geospatial data and imagery, enabling researchers to extract and analyze traffic sign information from real-world road environments.

3.3.1 Google Maps Street View

Google Maps Street View (GSV) is a widely used platform that provides panoramic imagery of streets across the globe (Figure 8). Its extensive coverage and accessibility make it an invaluable resource for traffic sign data collection. GSV imagery allows researchers to analyze traffic signs in their natural environments, offering insights into their placement, condition, and visibility. The platform supports geospatial mapping, enabling users to retrieve precise geographic coordinates for detected signs, which can then be integrated into traffic management systems.



Figure 8 Sample Image from Location (47.636855, -122.323283)

3.3.2 State Route Web Tool

The State Route Web (SRweb) Tool (<https://srview.wsdot.wa.gov/srweb/srweb.htm>) is a specialized application developed to provide a digital record of the Washington state highway system. It offers high-resolution roadway images and associated data, including milepost information and highway features, through an easy-to-navigate web interface (Figure 9). Users can explore specific regions and routes by selecting options from the SRweb map, which divides

Washington into six regions. Imagery is available for various years, with updates conducted biennially for half the system each year.

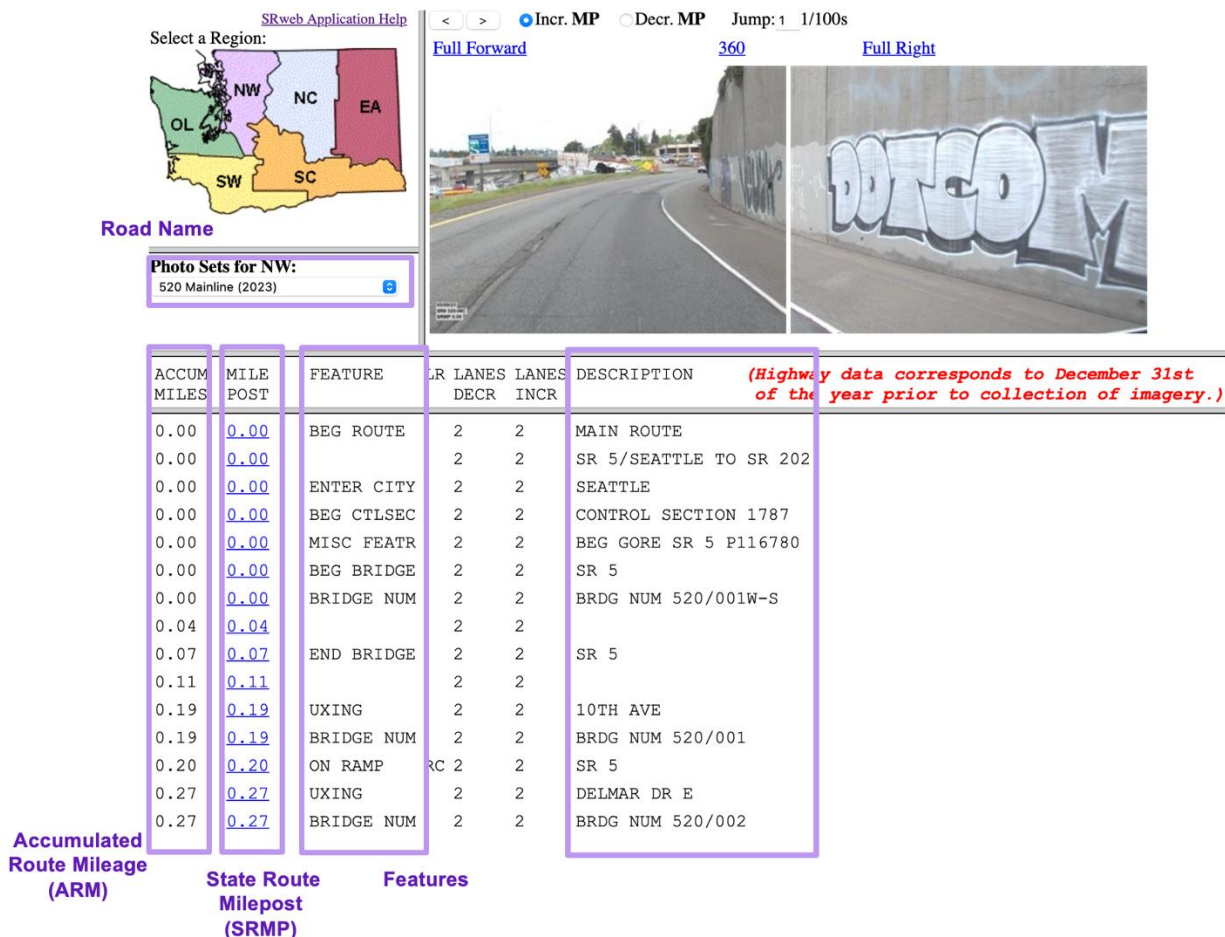


Figure 9 SRweb Tool Supported by WSDOT

SRweb's rich dataset is particularly useful for traffic sign studies, as it allows users to analyze traffic signs in situ, track changes over time, and extract valuable data for training and testing machine learning models. Its region-specific coverage and detailed annotations make it a practical resource for traffic sign research in Washington state.

3.4 DATASET SELECTION

Each of these datasets has unique strengths within the scope of this research but also presents some limitations. For example, the GTSRB is limited to the German traffic system,

making it less applicable to U.S. traffic environments. Similarly, the TT100K dataset does not include a sufficiently diverse representation of traffic signs from all geographic regions. The LISA dataset, despite being valuable for its focus on U.S. traffic signs, is relatively small in size and lacks the breadth needed to cover a wide range of real-world scenarios effectively. Finally, the images from the GSV dataset are not pre-annotated, necessitating manual annotation. These limitations created the need for additional data collection to ensure that a model could be tested and trained across a diverse array of traffic signs from various regions. Expanding the available dataset will allow models to adapt more effectively to a broader range of road conditions and traffic sign designs, promoting inclusivity and robustness in detection and recognition tasks.

To address these gaps, the research team divided the selected data sources into two groups and designed a two-model approach for this project. The first model was a vanilla traffic sign detection model, designed specifically for the data collection stage. This model focused solely on identifying the presence of traffic signs in images without requiring high classification accuracy. Its primary objective was to filter out frames that did not contain traffic signs, reducing the workload and transmission costs during the data collection process. By streamlining the data pipeline in this way, the team intended to ensure efficiency and minimize redundant work. The second model, a traffic sign detection and recognition model, was designed for the final system. With a more complex architecture, this model was tasked with accurately identifying and classifying traffic signs at a high precision level. Trained on carefully collected and labeled data, this model was intended to ensure robust and reliable performance in real-world scenarios, making it suitable for deployment in intelligent transportation systems.

To implement this approach, the team formalized two datasets tailored to the needs of each model. The first dataset, used for training the vanilla model, included three public

datasets: GTSRB, Tsinghua-Tencent 100K, and the LISA Dataset. While these datasets might not align perfectly with U.S. traffic conditions, they provided a foundational understanding of traffic sign features, such as shape and color, enabling the model to effectively filter relevant frames during data collection. **The second dataset, developed for the final detection and recognition model, incorporated the ARTS dataset, self-collected data, and open street images from Washington state.** This dataset was designed to ensure that the model was trained on diverse, representative data, including region-specific traffic signs, enabling it to perform reliably across varied road conditions. This two-dataset, two-model strategy ensured an efficient workflow while enhancing the robustness and accuracy of the final system. **Finally, images from the SRweb Tool were used to test the system's performance and build the sample inventory.**

4.0 TRAFFIC SIGN DATA COLLECTION AND DATASET CONSTRUCTION

4.1 OVERVIEW

The previous section reviewed and described the selection of several high-quality public traffic sign datasets to meet the requirements for model training and validation. While these datasets provide a strong foundation, they lacked the geographical diversity and real-world variability needed for robust model performance in Washington state.

To address these gaps, we designed a data collection system using dashboard cameras to capture traffic sign information across urban and rural areas in Washington state. This system collected diverse imagery under varying weather and lighting conditions, ensuring that the dataset captured real-world complexities. After manually reviewing and validating the labels, we integrated the self-collected data with the selected public datasets, forming a comprehensive and diverse dataset. This consolidated dataset served as a robust foundation to enhance the model’s generalization and adaptability, enabling it to achieve the highest possible performance and accuracy in real-world traffic environments.

4.2 VANILLA TRAFFIC SIGN DETECTION MODEL TRAINING

The main goal of this Traffic Sign Detection (TSD) model was to accurately differentiate and isolate traffic signs from their surrounding environment within an image during the data collection stage. However, because of the limited computing resources available on edge devices, we had to carefully consider the trade-off between precision and efficiency. To develop the most effective and efficient TSD model for our specific requirements, we comprehensively explored and evaluated various TSD model baselines. This thorough analysis ensured that we selected a model that would not only meet our precision needs but also operate effectively within the constraints of edge device computing resources. Those models included Faster R-CNN

Resnet 50, Faster R-CNN Resnet 101, Faster R-CNN Inception V2, Faster R-CNN Inception Resnet V2, R-FCN Resnet 101, SSD MobileNet V2, SSD Inception V2, and YOLO V4. A brief comparison of these models' advantages and shortcomings is shown in Table 2.

Table 2 TSD Model Comparison

Model	Advantages	Limitations
Faster R-CNN Resnet 50 and 101 (Ren et al. 2015)	The Faster R-CNN established itself as a predominant model in the field. It showcased a high accuracy in identifying and locating objects within images with great precision.	It exhibits substantial memory consumption and computational complexity, impeding the model's ability to deliver timely and efficient performance, which is often crucial in real-time object detection tasks.
Faster R-CNN Inception V2 and Inception Resnet V2 (Cai et al. 2018)	By leveraging the unique capabilities of the Inception architecture, the model can operate a variety of sizes of filters concurrently within each layer of the convolutional network, facilitating the extraction of a diverse range of features from the input data.	While the Faster R-CNN Inception V2 model offers a compelling balance between accuracy and computational efficiency, the model involves a complex architecture that requires extensive computational resources for training.
R-FCN Resnet 101 (Dai et al. 2016)	Attributable to the deep and complex Resnet 101 architecture, this R-FCN model enables the extraction of rich and hierarchical features from images, bringing accuracy and robustness in object detection.	R-FCN Resnet 101 demands significant computational resources and memory for both training and inference, which could be a limiting factor in resource-constrained environments.
SSD MobileNet V2 and Inception V2 (Liu et al. 2016)	SSD MobileNet V2 is famous for its lightweight architecture and speed, making it suitable for real-time processing and deployment on edge devices with limited computational resources.	While it offers a good balance between speed and accuracy, it might not match the precision levels of more complex models, potentially struggling with very small objects or objects with a high degree of occlusion.
YOLO V4 (Bochkovskiy et al. 2020)	YOLO V4 can maintain a stable frame rate of 30 FPS on a Jetson Nano device, demonstrating its potential for real-time, on-device traffic sign detection.	YOLO requires careful tuning and sufficient training data to optimize its performance and handle different challenging detection scenarios.

These selected models were trained on our first dataset, which consisted of selected images from three public traffic sign datasets. The performance of our model was evaluated on the basis of mean average precision (mAP), number of parameters, memory usage, and execution time, with results summarized in Table 3. Among the models, Faster R-CNN Inception ResNet V2 achieved the highest mAP (95.77), demonstrating superior performance. However, its high memory usage (18,250.45 MB) and execution time (442.22 ms) make it unsuitable for real-time

applications despite its accuracy. In contrast, SSD Mobilenet and SSD Inception V2, with lower mAP scores (61.64 and 66.10, respectively), exhibited reduced parameters and memory usage, making them better suited for resource-constrained systems. SSD Mobilenet had the shortest execution time, ideal for real-time processing. Models such as Faster R-CNN ResNet 101 and R-FCN ResNet 101 delivered high mAP scores (95.08 and 95.15, respectively), but they required significant computational resources, limiting their usefulness in real-time applications.

Table 3 Vanilla Traffic Sign Detection Model Comparison

Model	mAP	Parameters	Memory	Total_Exec_Millis
Faster R-CNN Resnet 50	91.52	43M	5256.454615	104.0363553
Faster R-CNN Resnet 101	95.08	62M	6134.705805	123.2729175
Faster R-CNN Inception V2	90.62	12M	2175.206857	58.53338971
Faster R-CNN Inception Resnet V2	95.77	59M	18250.446008	442.2206796
R-FCN Resnet 101	95.15	64M	3509.75153	85.45207971
SSD Mobilenet	61.64	55M	94.696119	15.14525
SSD Inception V2	66.10	13M	284.512918	23.74428378
YOLO V4	78.83	50M	1318.108256	21.4810122

Faster R-CNN Inception V2 achieved an accuracy of 90.62 percent (Figure 10), but its low memory consumption (2,175.20 MB) and modest parameter count (12M) made it suitable for edge deployment. Therefore, we deployed this model in the edge device and built a data collection system based on that. More details are discussed in the next section.



Figure 10 Sample Detection Results by the Faster-R-CNN-Inception-V2-Based TSD Model

4.3 DATA COLLECTION SYSTEM DEVELOPMENT

4.3.1 System Structure

The research team adopted a hands-on approach to data collection using self-equipped vehicles. These vehicles were used to capture high-resolution images of traffic signs under diverse weather conditions, varied lighting, and across a range of traffic scenarios. The vehicles were equipped with advanced technologies, including dash cameras (specifications provided in Table 4), edge-computing devices, and communication modules. These tools were deployed strategically along highways and local routes, enabling the collection of raw video and image data of traffic signs.

Table 4 Relevant Specifications of the Data Collection Vehicle Dash Camera

Specification	Detail
Resolution	1920×1080P @ 30fps
Field of View	150° A+ Ultra HD Wide Angle
Light Sensor	5.0 PM
Lens Aperture	F 1.8

The data flow of the self-data collection system is shown in Figure 11 and Figure 12. Dash cameras served as the primary data source, capturing raw imagery. The raw data were then transmitted to onboard edge-computing devices using a local network. These edge devices were

equipped with the vanilla traffic sign detection model we had trained before, which processed the raw input to extract traffic sign data and delete the relevant frames. The final output consisted of isolated traffic sign image regions and their initial classification results. These results were transferred to a back-server at the Smart Transportation Applications and Research Laboratory (STAR Lab) at the University of Washington for further analysis and data management. This comprehensive approach allowed the team to collect practical, real-world data under various conditions, enhancing the robustness and accuracy of the Traffic Sign Detection and Recognition model.

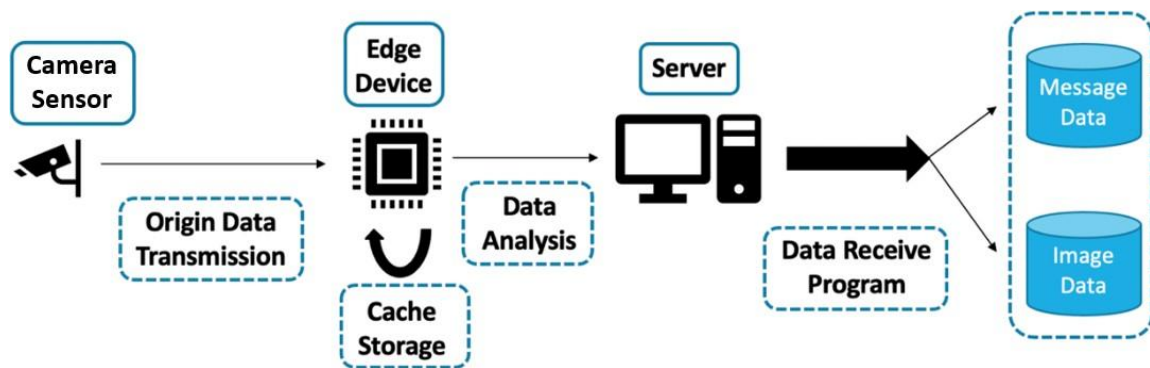


Figure 11 The Structure of Data Collection System

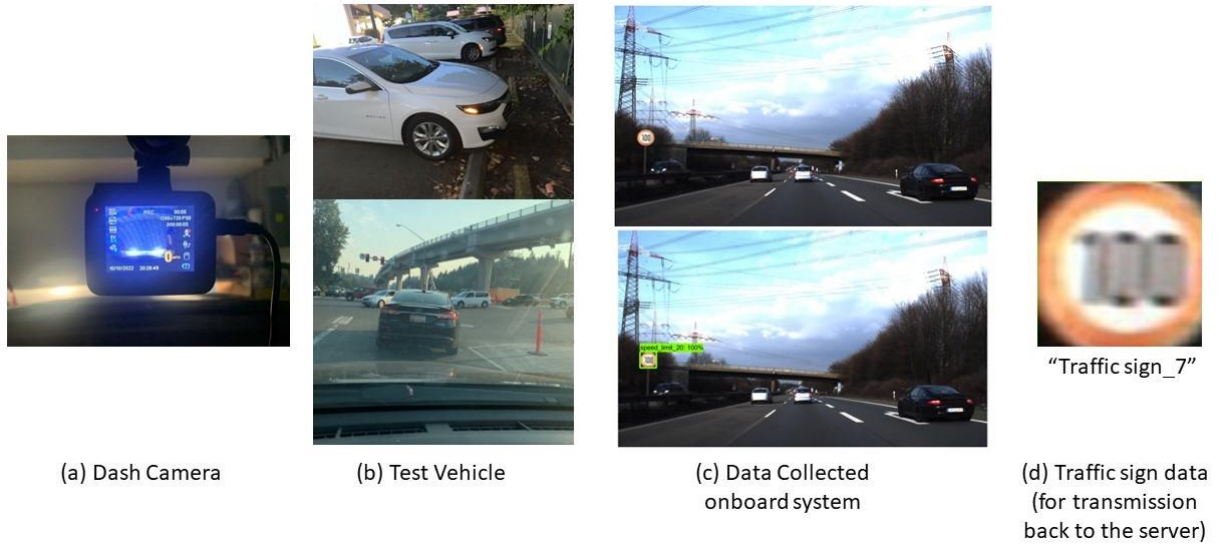


Figure 12 Collection System Overview

4.3.2 Data Collection Process

To develop a robust Traffic Sign Detection and Recognition model, the research team conducted an extensive data collection process using self-equipped vehicles (see Figure 13). Over two hours of video footage was recorded at a frame rate of 30 FPS, resulting in a comprehensive dataset of traffic scenes. From this raw footage, more than 2,000 traffic signs were manually selected and labeled, creating a high-quality annotated dataset for training and validation purposes. The collection included traffic signs from both highways and local roads, ensuring a diverse range of data to reflect real-world traffic conditions.

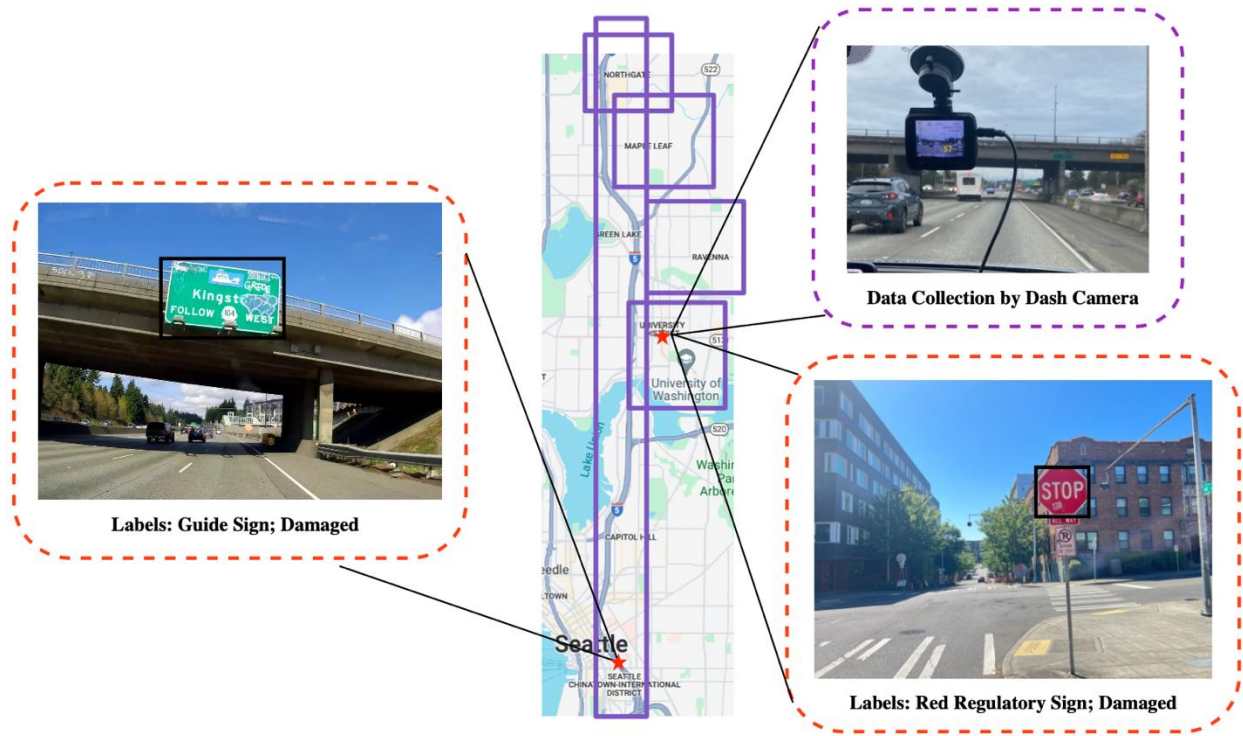


Figure 13 Data Collection Process

The data were collected across various locations, including the University District, Ravenna, Maple Leaf, and Northgate areas of Seattle, as well as along the I-5 freeway, covering both urban and freeway settings. To ensure that the dataset captured a wide range of conditions, the collection took place during noon and afternoon hours under both sunny and cloudy weather. This diversity in lighting and environmental factors enhanced the dataset's versatility, enabling the model to generalize effectively to varied scenarios. By incorporating data from diverse areas and conditions, this process ensured that the dataset provided a solid foundation for building a reliable and adaptable Traffic Sign Detection and Recognition model.

4.3.3 Post-Processing

To ensure accurate annotations of traffic signs in each image, a comprehensive post-processing procedure was implemented. This process involved selecting specific frames from the saved video footage, cropping them to an appropriate size, and labeling the traffic signs

according to their categories. This step was critical to maintaining consistency with images from other selected datasets while mitigating potential discrepancies caused by variations in collection methods and sources. By standardizing the annotation process, the research team aimed to enhance the reliability and robustness of the resulting dataset.

Moreover, the vanilla traffic sign detection model initially categorized traffic signs into 43 classes. However, these categories didn't include all traffic sign types relevant to our application scenarios, and such classification cannot benefit the maintenance procedure processed by government and agencies. **To address this, the labels were simplified into seven major categories based on their color patterns and aligned with the minimum retro-reflectivity levels outlined in the Manual on Uniform Traffic Control Devices (MUTCD). These categories comprised Guide Signs (Information), Guide Signs (Route), Red Regulatory Signs, Black and White Regulatory Signs, Warning Signs, School Signs, and Markers.** This simplified classification approach not only broadened the coverage of traffic sign types but also enhanced the model's ability to detect damaged signs, as highly detailed classifications often lead to missed detections due to the loss of key features. By categorizing traffic signs on the basis of their functions and color patterns, this method ensured system performance and robustness while supporting the additional focus on traffic sign condition assessment. Figure 14 shows the labeling process.

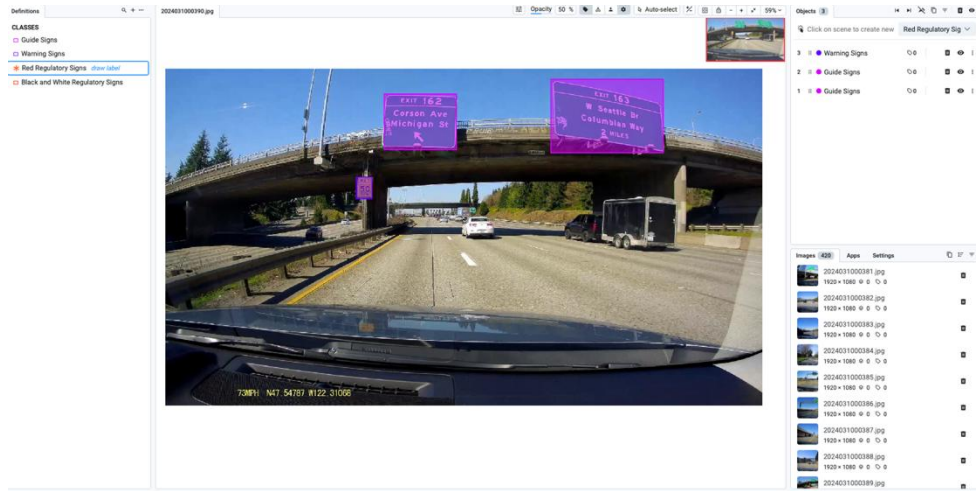


Figure 14 Labeling Process

4.4 DATASET CONSTRUCTION

The final dataset used for this project consisted of **8,057 images** from three primary sources. Specifically, **4,000 images** were drawn from the ARTS dataset, **853 images** were used from open street imagery, and the remaining **3,204 images** were obtained through our self-collection efforts. This diverse combination ensured a comprehensive dataset capable of supporting robust model training and validation.

The dataset's traffic signs were classified into seven main categories: Guide Signs (Information), Guide Signs (Route), Red Regulatory Signs, Black and White Regulatory Signs, Warning Signs, School Signs, and Markers. The class distribution was imbalanced, with Guide Signs (Information) and Regulatory Signs (Black and White) being the most prevalent (Figure 15). Although this reflected the natural occurrence of traffic signs, additional techniques such as data augmentation or class balancing might be applied during training to improve model performance across underrepresented classes.

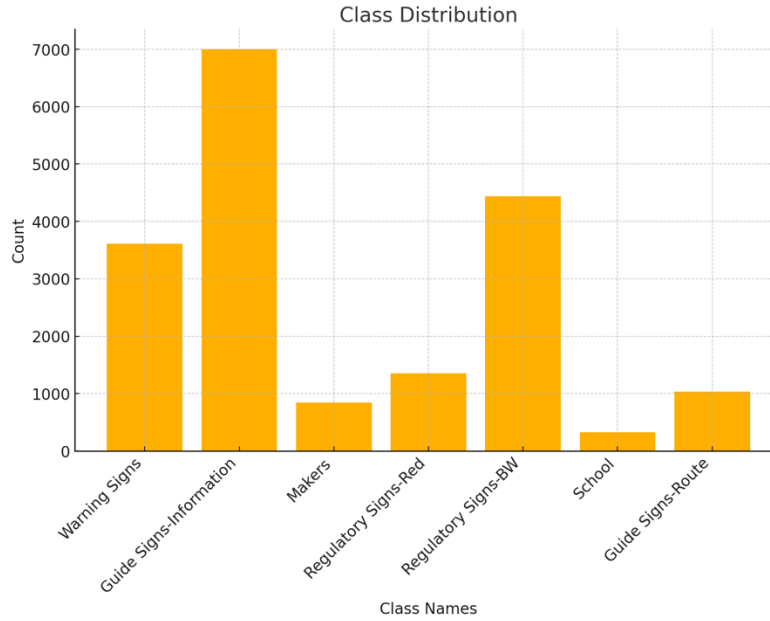


Figure 15 Class Distribution

The distribution of objects per image (Figure 16) revealed that most images contained fewer than five traffic signs, reflecting the typical sparsity of signs along real-world roadways. This pattern helped ensure that the model was trained to perform well in both densely and sparsely populated scenarios. Additionally, the spatial distribution of annotated signs, as indicated by the heatmap analysis, showed that traffic signs were primarily concentrated in specific areas of the images, corresponding to their usual placement along roadsides. This feature allowed the model to focus on relevant regions while ignoring irrelevant background details.

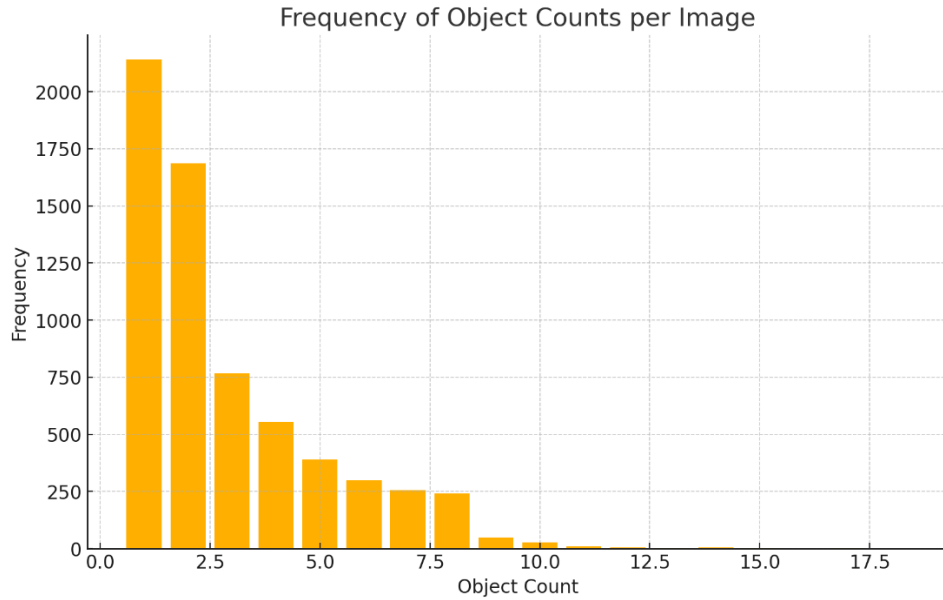


Figure 16 Object Count Per Image

Each data source in the dataset added a unique value. The ARTS dataset provided high-quality annotated images (Figure 17), serving as a solid baseline for training. Open street images contributed geographic diversity and real-world variability, complementing the controlled conditions of ARTS data. Meanwhile, self-collected data, gathered from urban and rural areas in Washington state under various lighting and weather conditions, further enhanced the dataset's adaptability to real-world scenarios. Together, these sources created a well-rounded and representative dataset, enabling the development of a robust Traffic Sign Detection and Recognition model.

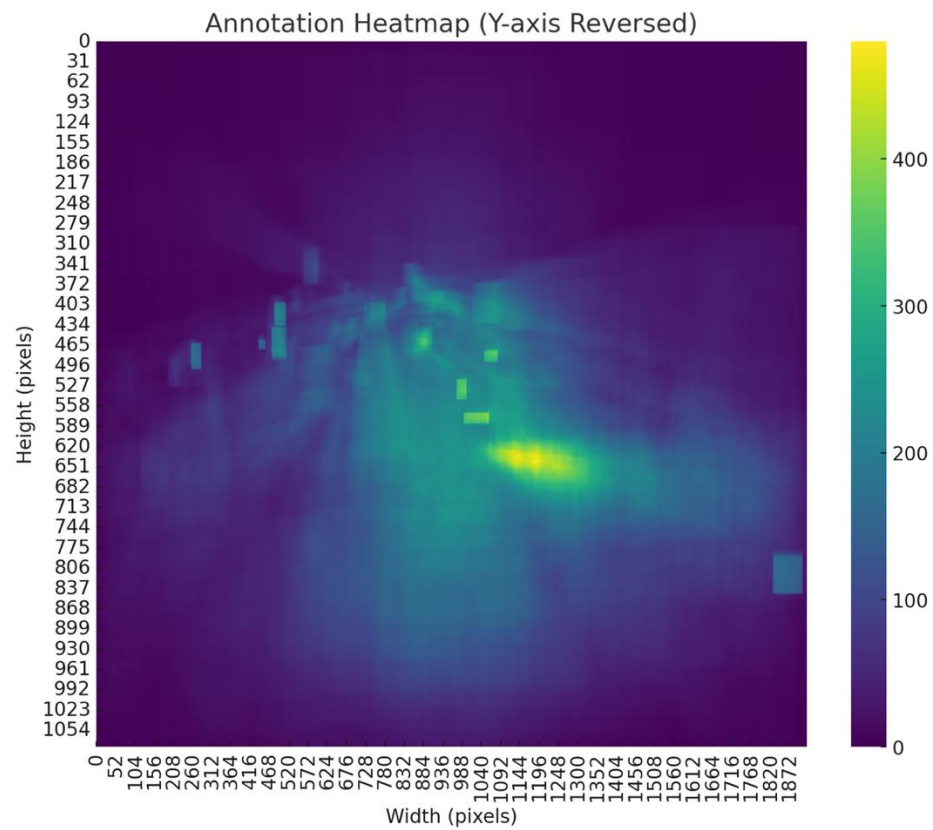


Figure 17 Annotation Heatmap

5.0 LEARNING-BASED TRAFFIC SIGN DETECTION AND RECOGNITION ALGORITHM DEVELOPMENT

5.1 OVERVIEW

The development of a learning-based traffic sign detection and recognition algorithm was a critical component of this research. This chapter outlines the systematic approach we used to select, modify, train, and evaluate models to achieve optimal performance in traffic sign detection and classification. It begins with a review of five state-of-the-art object detection models, assessing their strengths and weaknesses to identify the most suitable candidate. Sparse R-CNN, known for its robust accuracy and efficiency, was selected as the final model, with modifications made to adapt it to the specific requirements of this project.

The training process leveraged a high-performance computing environment and carefully calibrated hyperparameters to enhance model performance under diverse conditions. A combination of advanced loss functions ensured that the model achieved precise localization, accurate classification, and robust generalization across various traffic sign types. Performance evaluation was conducted using industry-standard metrics such as mAP, memory usage, and frames per second (FPS), offering a comprehensive understanding of the model's performance.

5.2 MODEL SELECTION AND MODIFICATION

For developing a robust traffic sign detection and recognition system, selecting appropriate models was crucial to balancing accuracy and speed based on application needs. For this research, five models were evaluated: **SSD-VGG16-512**, **Faster-RCNN-R101-FPN**, **YOLOX-l**, **Deformable DETR-R50**, and **Sparse R-CNN**. Each model was chosen for its distinct approach to object detection, representing a range of trade-offs between speed and accuracy. After careful analysis of their performance and suitability for real-world scenarios,

Sparse R-CNN was ultimately selected as the final model for its superior balance of performance and practicality.

The **SSD-VGG16-512** and **Faster-RCNN-R101-FPN** models were initially used for vanilla model training. However, as accuracy became a higher priority, we switched to more complex backbones and retrained them on our new dataset.

The **YOLOX-l** model introduces an anchor-free approach to object detection, simplifying the architecture and improving efficiency (Ge et al. 2021). By decoupling classification and regression heads, YOLOX achieves faster inference times without sacrificing much accuracy. Its lightweight design and real-time performance make it a strong contender for applications that require high-speed detection. However, while YOLOX excels in efficiency, it does not achieve the same level of precision as more complex models like Sparse R-CNN or Deformable DETR.

The **Deformable DETR-R50** combines transformer-based object detection with deformable attention mechanisms (Zhu et al. 2021). This model excels in accurately detecting objects in challenging scenarios, such as occlusions or varying scales, by focusing on relevant regions of interest while reducing computational overhead. However, while Deformable DETR offers exceptional precision, particularly for complex scenes, its slower inference speed and high computational requirements make it less suitable for time-critical applications.

The **Sparse R-CNN** model was ultimately chosen for its ability to strike a balance between accuracy and efficiency (Sun et al. 2021). By using sparse object queries instead of dense region proposals, Sparse R-CNN significantly reduces computational costs while maintaining high detection accuracy (Figure 18). Its simplified structure and robust performance in detecting and classifying traffic signs, even in challenging conditions, made it the most

practical option for this project. Sparse R-CNN outperformed SSD and YOLOX in accuracy while being more efficient than Deformable DETR and Faster R-CNN.

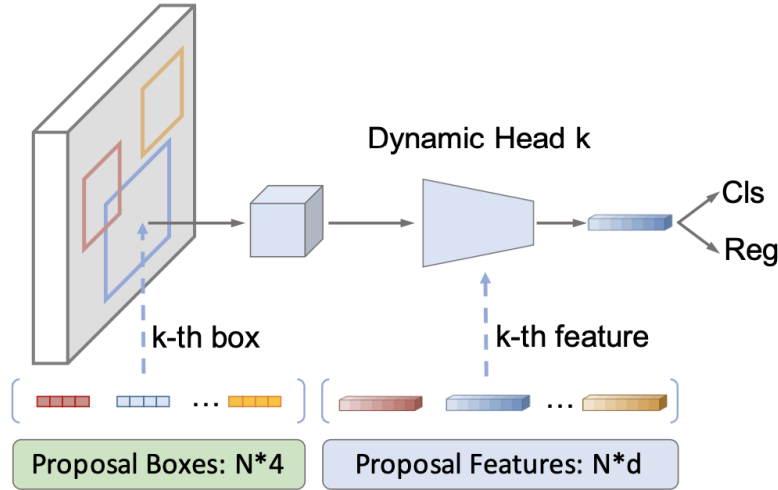


Figure 18 Sparse RCNN Architecture (Sun et al. 2021)

To deploy the Sparse R-CNN model effectively for traffic sign detection and recognition, several modifications were made to adapt it to the project's specific requirements. The model was configured to use ResNet-101 as the backbone, leveraging its strong feature extraction capabilities for handling diverse traffic sign characteristics. Additionally, the classification heads were customized to align with the output requirements, ensuring compatibility with the simplified traffic sign categories used in the dataset. These adjustments enhanced the model's performance by improving its ability to detect and classify traffic signs accurately while maintaining computational efficiency, making it well-suited for real-world deployment scenarios.

In conclusion, each of the evaluated models brought unique strengths and trade-offs for this task. While SSD and YOLOX excelled in speed and efficiency, their accuracy fell short for certain traffic sign types. Deformable DETR and Faster R-CNN offered superior precision but

were less practical for real-time applications. Sparse R-CNN, with its robust performance and balanced efficiency, was selected as the final model. The modifications made for deployment further enhanced its effectiveness, ensure that it met the project’s goals for accurate and reliable traffic sign detection and recognition.

5.3 TRAINING PROCESS

5.3.1 Environment Preparation

The hardware utilized for this research was a high-performance workstation located in the STAR Lab. The system featured 94.0 GiB of random-access memory (RAM), providing ample memory to process large datasets and manage training operations efficiently. It was powered by an Intel® Core™ i9-7900X central processing unit (CPU) with 20 cores running at 3.30 GHz, which offered significant computational power for multi-threaded tasks. Additionally, the system was equipped with two NVIDIA GeForce GTX 1080 Ti graphic processing units (GPUs), enabling accelerated parallel processing through Compute Unified Device Architecture (CUDA) support.

On the software side, the operating system was Ubuntu 20.04. The primary deep learning framework employed was PyTorch 2.1.1, and CUDA Toolkit 12.0.1 was used to leverage GPU acceleration, significantly reducing training times and improving overall efficiency. The MMDetection library, version 3.3.0, was integrated as the core object detection toolbox. MMDetection provided pre-implemented Sparse R-CNN architectures and customizable modules, streamlining the development and experimentation processes.

5.3.2 Training Strategies

The training process for the Sparse R-CNN model was designed to optimize performance across a range of challenging conditions while maintaining computational efficiency.

Hyperparameters

A ResNet-101 backbone was employed and pre-trained on the ImageNet dataset to ensure robust initialization. For the training process, the model was trained end-to-end for a total of 54 epochs, with milestones at epochs 38 and 48 reducing the learning rate by a factor of 0.1.

Key hyperparameters included a learning rate of $2.5e-5$, optimized by using the AdamW optimizer with a weight decay of $1e-4$. A warm-up learning rate strategy was applied over the initial 500 iterations to ensure smooth convergence. The batch size was set to 16, balancing training speed and hardware resource limitations. To improve generalization, data augmentation techniques were implemented, including random scaling, flipping, cropping, and resizing to simulate real-world variations in traffic sign appearance. Multi-scale training was employed, resizing images to scales ranging from 400 to 800 pixels to improve the model's ability to detect signs at varying distances.

Loss Functions

We employed three key loss functions—Focal Loss, L1 Loss, and Generalized Intersection over Union (GIoU) Loss—to optimize its performance for traffic sign detection and recognition (Sun et al. 2021).

The Focal Loss is designed to address the problem of class imbalance in object detection tasks, where the number of negative samples (background regions) significantly outweighs the number of positive samples (traffic sign regions). Traditional loss functions like Cross-Entropy Loss tend to be dominated by easy-to-classify examples, leading to suboptimal performance on harder, minority-class samples. Focal Loss mitigates this by down-weighting the loss contribution of well-classified examples and focusing the model's learning on harder-to-classify examples. The function is as follows:

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t)$$

where p_t is the predicted probability of the true class. α_t balances the importance of positive and negative samples (set to 0.25 in this project). γ controls the rate at which the loss contribution from well-classified examples decreases (set to 2.0 here).

The L1 Loss, also known as the mean absolute error (MAE), is used for bounding box regression, ensuring that the predicted bounding boxes closely match the ground truth. This loss function measures the absolute differences between the predicted bounding box coordinates and their ground truth values. L1 Loss is particularly effective for achieving the precise localization of traffic signs, as it penalizes even small deviations. The expression for L1 Loss is as follows:

$$L_{L1} = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i|$$

where N is the total number of bounding boxes. x_i and $\hat{x}_i$ represent the ground truth and predicted bounding box coordinates, respectively.

The GIoU Loss improves upon traditional IoU-based losses by addressing scenarios in which the predicted and ground truth bounding boxes do not overlap. While IoU Loss only penalizes the intersection-to-union ratio, GIoU Loss also considers the smallest enclosing box that covers both the predicted and ground truth boxes, encouraging better alignment and coverage. The expression for GIoU Loss is as follows:

$$L_{GIoU} = 1 - \left(IoU - \frac{A_c - A_u}{A_c} \right)$$

where IoU is the Intersection over Union between the predicted and ground truth boxes. A_u is the area of the union of the two boxes. A_c is the area of the smallest enclosing box.

During the training stage, L1 Loss was assigned a weight of 5.0 to reflect the critical importance of accurate localization in traffic sign detection. GIoU Loss was assigned a weight of

2.0 to complement the L1 Loss by penalizing shape and position mismatches, particularly in cases where there was little or no overlap between the bounding boxes. This ensured that the model not only achieved precise localization but also improved the spatial alignment of its predictions. Figure 19 shows the three different loss changes during the training process.

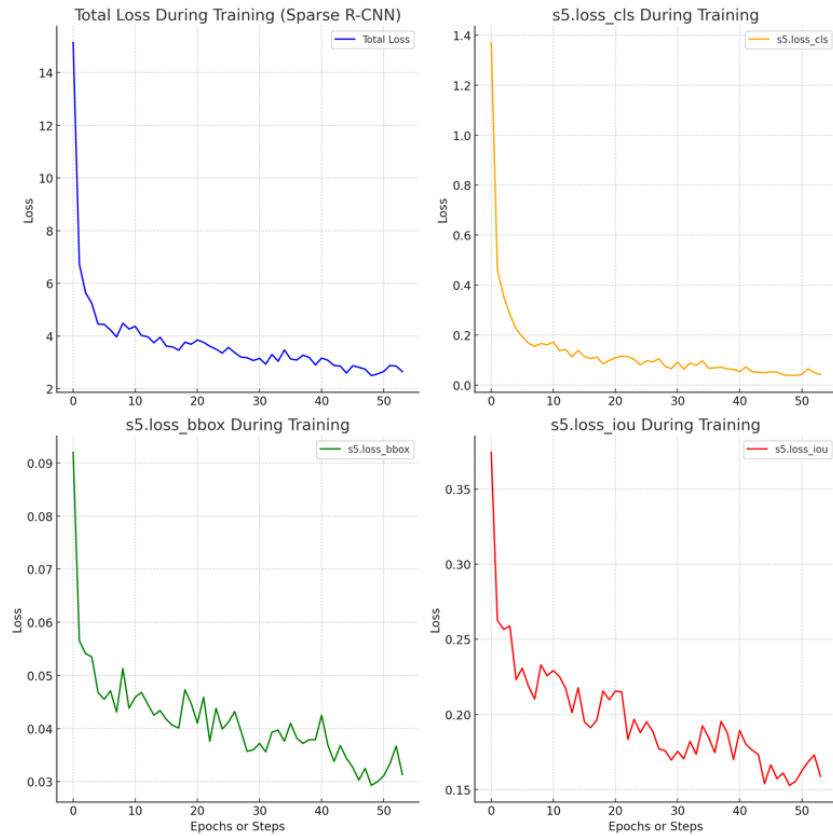


Figure 19 Sparse R-CNN Loss Changes during Training

5.4 PERFORMANCE COMPARISON

5.4.1 Evaluation Metrics

To evaluate the performance of the selected object detection models, several key metrics were utilized, offering insights into both accuracy and computational efficiency. A central metric for accuracy was mAP, which measured the average precision across multiple Intersection over Union (IoU) thresholds. The formula for mAP is as follows:

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i$$

where N represents the total number of object classes, and AP_i is the Average Precision for class i . Average Precision itself is derived from the Precision-Recall curve, where Precision quantifies how many of the predicted positive instances are correctly labeled. Precision can be expressed as follows:

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

Specific mAP thresholds, such as mAP_{50} (IoU = 0.5) and mAP_{75} (IoU = 0.75), assessed performance under varying levels of strictness. Additionally, metrics such as mAP_s , mAP_m , and mAP_l evaluated model accuracy on small, medium, and large objects, respectively, ensuring that size-dependent performance was accounted for. Efficiency metrics included memory usage (Mem), which measured GPU memory consumption during inference, and FPS, which reflected the model’s real-time processing capability. Together, these metrics provided a comprehensive evaluation of each model’s trade-offs between accuracy and computational resource requirements.

5.4.2 Results

Performance Comparison of Object Detection Models

The performance of Sparse R-CNN was compared with that of four other mentioned object detection models using our dataset. The results are summarized in Table 5. We employed the default training parameters for each model as outlined in their original studies, as our primary objective was not to optimize individual model performance but rather to evaluate their comparative effectiveness. **Sparse R-CNN consistently achieved the highest performance across nearly all metrics, particularly excelling with the mAP_{50} standard, under which it**

reached an impressive 93.5 percent. This highlighted its robustness in detecting traffic signs under less stringent conditions, a critical feature for real-world applications in which a perfect match with ground truth data is not always necessary. Additionally, Sparse R-CNN demonstrated superior performance under the more rigorous mAP_{75} standard, achieving a score of 90.4 percent. This indicated that Sparse R-CNN not only detects a greater number of traffic signs but also ensures higher precision in region-of-interest (ROI) selection in comparison to other models.

Table 5 Performance Comparison

Model	mAP	mAP₅₀	mAP₇₅	mAP_s	mAP_m	mAP_l	Mem(GB)	FPS
SSD-VGG16-512	68.20	86.50	80.00	31.90	60.20	76.20	5.65	62.9
Faster-RCNN-R101-FPN	75.50	91.60	87.10	46.00	70.90	81.40	6.25	18.4
YOLOX-l	68.00	83.40	76.80	24.00	58.80	76.80	6.16	91.5
Deformable DETR-R50	77.60	92.70	89.40	47.20	70.20	85.20	6.29	22.4
Sparse-RCNN R-101-FPN-300	79.20	93.50	90.40	45.00	74.10	85.60	6.20	25.4

One of the most significant challenges in traffic sign detection is the ability to accurately identify small objects, as these signs often appear at a distance in real-world applications. The mAP_s scores across all models indicated that small object detection remains a difficult task. However, Sparse R-CNN achieved the highest mAP_s score at 45.0 percent, surpassing other models in its ability to detect small traffic signs. For medium-sized traffic signs (mAP_m), Sparse R-CNN achieved 74.1 percent, which was slightly lower than Deformable DETR's 70.2 percent but still among the highest-performing models. In terms of large object detection (mAP_l), Sparse R-CNN again led the models with 85.6 percent, showcasing its capacity to handle both small and large traffic signs effectively.

Training Performance and Convergence Analysis

The mAP_{50} changes during training, as shown in Figure 20, provided insights into how each model's performance evolved over time. Sparse R-CNN displayed a stable learning curve, rapidly improving in the early training epochs and maintaining a consistently high mAP_{50} score throughout training. For these models, DETR and Faster R-CNN exhibited the most stable and high-performing training curves, consistently achieving mAP_{50} values above 90 percent after approximately 15 epochs. Sparse R-CNN closely followed, with a slightly higher final performance than DETR, demonstrating its adaptability and effectiveness in learning sign features. SSD and YOLOX, while faster models, exhibited slower convergence and lower overall mAP_{50} scores. SSD plateaued at a lower score, indicating its limitations in achieving high accuracy despite its speed advantage. YOLOX demonstrated an initial rapid improvement but plateaued below the performance of the other high-accuracy models, suggesting its reliance on speed over detection precision.



Figure 20 mAP_{50} Changes during Training

These observations confirmed that while some models can achieve high accuracy, they require more training epochs to stabilize, whereas Sparse R-CNN consistently achieves high performance without excessive computational overhead.

Resource Efficiency and Speed Analysis

From a resource perspective, Sparse R-CNN required only 6.2 GB of memory, which was less than other high-performance models such as Faster R-CNN and Deformable DETR, and comparable to faster models such as YOLOX and SSD that prioritize speed. This moderate memory usage enables Sparse R-CNN to maintain high detection accuracy without imposing excessive demands on system resources.

In terms of processing speed, Sparse R-CNN achieved 25.4 FPS, outperforming other high-accuracy models while trailing speed-optimized models such as YOLOX and SSD.

Although not the fastest, Sparse R-CNN's FPS rate was sufficient to meet the requirements of our system. In summary, Sparse R-CNN offers an excellent balance between speed, accuracy, and resource efficiency, making it an ideal choice for our application.

Detection Example Analysis

Figure 21 presents several real-world detection examples captured from a dashcam perspective that demonstrate the model's ability to accurately identify traffic signs in diverse driving environments. The images showcase traffic signs at varying distances, with some appearing close to the camera and others farther away. Despite these differences in scale, the model successfully detected both large overhead highway signs and smaller roadside signs, highlighting its robustness in handling traffic signs of different sizes.

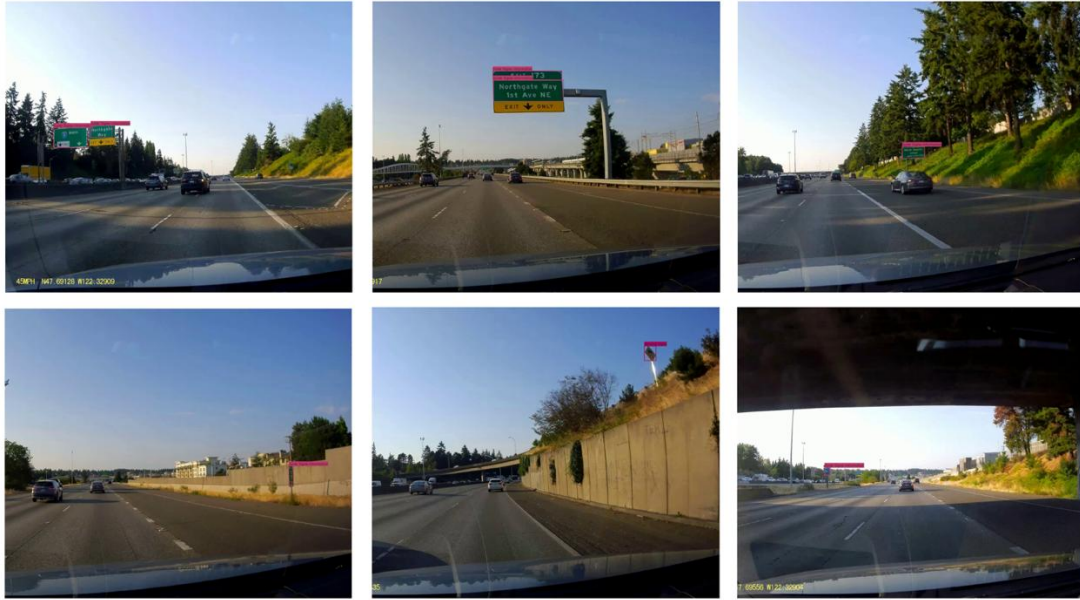


Figure 21 Some Detection Examples

Additionally, the model performed well under various lighting conditions, including bright daylight and shadows. Even when signs were partially shaded or set against bright skies, the detection remained consistent, indicating the model's resilience to illumination challenges. Another key observation was the model's ability to detect signs in complex environments where occlusions occurred. Some signs were partially obstructed by trees, highway barriers, or poles, yet the model successfully identified them, demonstrating strong feature extraction capabilities. Overall, these examples illustrate the model's strong performance in real-world conditions, making it suitable for practical applications in traffic sign detection and recognition.

6.0 TRAFFIC SIGN CONDITION ASSESSMENT

6.1 OVERVIEW

In this project, one of our goals was to automate the traffic sign assessment process by leveraging advanced CV and machine learning techniques. In this chapter, we propose a condition assessment framework that uses cropped traffic sign images as input to assess their conditions as either Healthy or Damaged. By utilizing six retro-reflectivity-related features, our model provides a lightweight, efficient, and accurate method for assessing traffic sign conditions. The chapter outlines the methodologies for feature selection and extraction, the design of a deep neural network (DNN) model for classification, and the experimental validation of this system. The proposed approach demonstrated higher accuracy and computational efficiency than conventional, full-image-based methods.

6.2 METHODOLOGIES

Traditional image-based assessment methods often use entire images as input to classify conditions. While these high-dimensional data provide comprehensive information, they also significantly increase computational costs by requiring the processing of unnecessary features. In practical applications, deploying another complex, deep learning model to assess traffic sign conditions after detection would drastically reduce processing speed and hinder the transition of the system to edge devices. To address this, we opted to manually select highly relevant features and to employ a simpler machine-learning model. This approach streamlines the system, enhances processing speed, and maintains efficiency.

6.2.1 Imagery Assessment Features Selection and Extraction

The selection of features for traffic sign condition assessment was guided by their relevance to retro-reflectivity and their ability to capture damage characteristics effectively.

Traditional methods relying on entire image analysis often fail to account for subtle damage patterns resulting from background interference and exposure variations. To overcome these challenges, six features were identified as critical for assessing traffic sign conditions: **texture**, **shape**, **saturation**, **brightness**, **contrast**, and **category**. These features were selected because of their ability to efficiently represent the physical state of traffic signs, as described below.

Texture

The surface texture of a traffic sign is closely associated with its retro-reflectivity, as uneven surfaces or fading can significantly affect its visibility. By employing the **Gray-Level Co-Occurrence Matrix (GLCM)** (Haralick et al. 1973), the texture is quantified through five metrics: contrast, dissimilarity, homogeneity, energy, and correlation. These metrics capture the uniformity, variability, and spatial relationships of pixel intensities, making them highly effective for identifying indicators of wear or fading. For example, high contrast or low homogeneity values often indicate surface degradation. The expressions of each metric are as follows:

$$\begin{aligned}
 \text{Contrast} &= \sum_{i,j=0} P(i,j) \cdot (i - j)^2, \\
 \text{Dissimilarity} &= \sum_{i,j=0} P(i,j) \cdot |i - j|, \\
 \text{Homogeneity} &= \sum_{i,j=0} \frac{P(i,j)}{1 + (i - j)^2}, \\
 \text{Energy} &= \sum_{i,j=0} P(i,j)^2, \\
 \text{Correlation} &= \sum_{i,j=0} \frac{(i - \mu_i)(j - \mu_j) \cdot P(i,j)}{\sigma_i \sigma_j}
 \end{aligned}$$

where $P(i, j)$ represents the normalized value at position (i, j) in GLCM.

Shape

Damages such as bending, tearing, and deformation alter the original shape and size of traffic signs. To capture these changes, **Hu moments** were chosen for their invariance to translation, rotation, and scaling (Hu et al. 1962). These seven invariant moments, shown below, provide a robust mathematical representation of shape-related features, enabling the detection of structural anomalies without being affected by the sign's orientation or position.

$$\begin{aligned}\phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - \\ &\quad (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]\end{aligned}$$

Saturation

The color saturation of a traffic sign is a key indicator of its retro-reflectivity and condition. For example, as shown in Figure 22, higher saturation values typically correlate with newer, well-maintained signs, whereas older or damaged signs exhibit more uniform and lower saturation distributions. By analyzing the variance of saturation values, this feature captures both local and global changes in color intensity, making it an effective marker for identifying discoloration or fading.

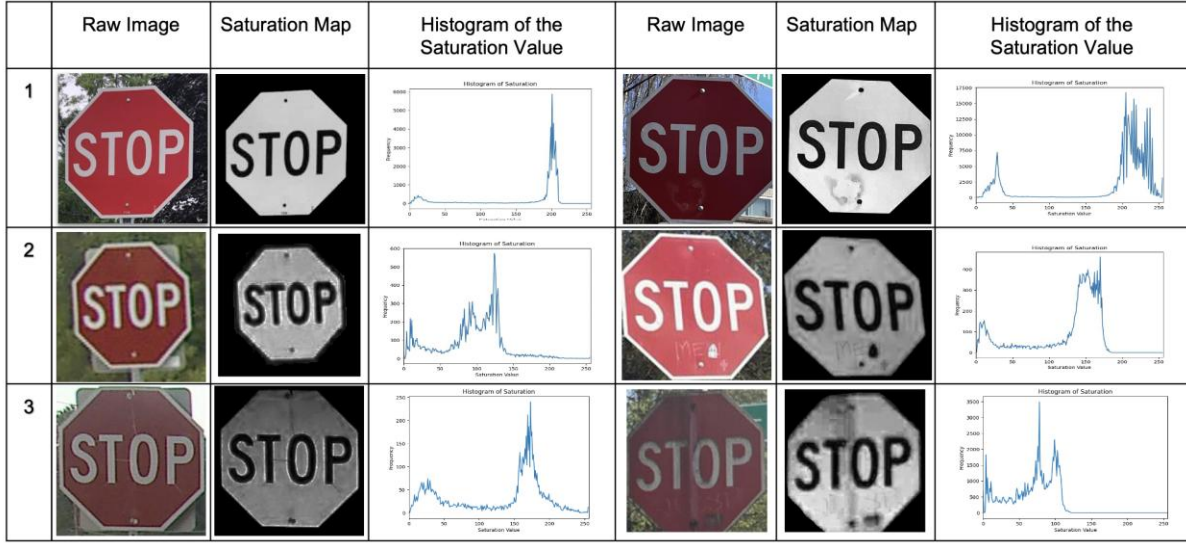


Figure 22 The Saturation Maps and Histogram of Saturation Value of Stop Signs in Different Ages

Brightness

Brightness reflects the overall light intensity of a traffic sign and is directly influenced by its retroreflective performance. Dim or uneven brightness often indicates wear, fading, or damage. Including brightness as a feature ensures that the model considers the sign's visibility under different lighting conditions. The brightness B is defined as follows:

$$B = \frac{1}{N} \sum_{i=1}^N I_i$$

where N represents the total number of pixels in the sign's region, and I_i denotes the intensity of each pixel. This formulation provides an average measure of the sign's brightness, aiding in the detection of deteriorated or poorly visible traffic signs.

Contrast

Contrast is critical for assessing the clarity of a traffic sign's text and symbols, which are essential for readability. Low contrast often suggests wear or fading, reducing the sign's effectiveness. This feature was selected to capture variations in visual clarity, ensuring that the

model accounts for changes that may affect the sign's functionality. The contrast of a traffic sign is calculated as follows:

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

where I_{max} and I_{min} are the maximum and minimum intensity values in the traffic sign area, respectively. Figure 23 shows brightness and contrast heatmaps.



Figure 23 Brightness and Contrast Heatmaps for Traffic Signs

Category

In our detection model, traffic signs are classified into seven main categories based on their predominant color patterns. This classification aligns with the MUTCD guidelines and simplifies the assessment process by grouping signs with similar retroreflective properties. Including the category as a feature allows the model to tailor its assessment based on the specific characteristics of each group. Figure 24 shows the seven traffic sign categories based on their predominant color patterns: White on Green, White on Blue, White on Brown, Black on Yellow, White on Red, and Black on White. The retroreflectivity values, measured in candela per lux per

square meter (cd/lx/m^2) under specific observation and entrance angles, help determine whether a sign meets the required visibility standards.



Minimum Maintained Retroreflectivity Levels (cd/lx/m^2) for six kinds of traffic sign color patterns in MUTCD
(The levels are for prismatic sheeting traffic signs and are measured at an observation angle of 0.2° and an entrance angle of -4.0°)

Figure 24 Minimum Maintained Retro-reflectivity Levels in MUTCD

In our classification system, each traffic sign color pattern corresponds to one or two specific categories to better align with their functional roles. The White on Green, White on Blue, and White on Brown signs primarily fall under the Guide and Information Signs and Route Marker Signs, which provide navigation and roadway information. White on Red signs belong to the Red Regulatory Sign category, typically indicating stop, yield, or prohibition-related regulations. Black on White signs are classified under Black and White Regulatory Signs, which include speed limits, lane usage, and other enforceable traffic rules. Black on Yellow corresponds to the Warning Sign category, which alerts drivers to upcoming road conditions, hazards, or changes in roadway geometry. Additionally, specific signs with fluorescent yellow-green backgrounds, often used in school zones and pedestrian crossings, fall under the School Sign category. By structuring traffic sign classification this way, our model ensures a more precise assessment, allowing for targeted evaluation based on the retroreflective and functional characteristics of each sign type.

6.2.2 Model Selection

To classify traffic signs on the basis of their condition, a lightweight **deep neural network (DNN)** was implemented. Unlike complex convolutional networks, this six-layer DNN relies on the manually extracted features, reducing computational complexity while maintaining high accuracy. The DNN begins with an input layer for the six features, followed by fully connected layers with progressively fewer units. Each layer includes **batch normalization**, **ReLU activation**, and **dropout layers** to prevent overfitting and to improve training efficiency. The compact architecture (Figure 25) ensures that the model is efficient enough for deployment on edge devices, making it suitable for real-time traffic sign condition assessment.

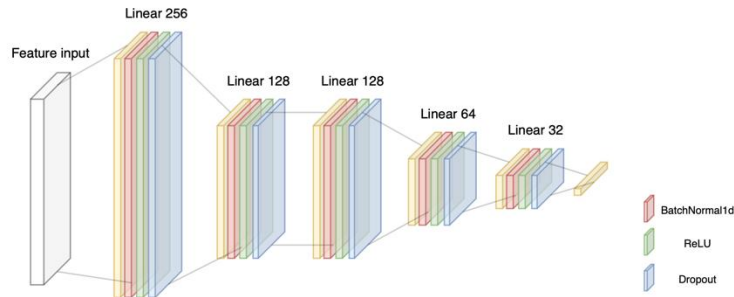


Figure 25 Assessment Model Architecture

6.3 EXPERIMENTS

The performance of the proposed traffic sign condition assessment models was evaluated with several metrics, including **accuracy**, **precision (P)**, **recall (R)**, **F1-score**, **model parameters (Para)**, and **inference time (Infer)**. These metrics were analyzed for two classes: **healthy signs (Class 0)** and **damaged signs (Class 1)**. Generally, the results highlighted the competitive performance of our proposed method in comparison to that of the YOLO models, but it required many fewer parameters.

6.3.1 Methods Comparison

The DNN classifier based on selected features demonstrated the best overall performance, achieving an accuracy of 93.56 percent and outperforming both the YOLOv8-n and YOLOv8-l models, which achieved accuracies of 92.04 percent and 92.79 percent, respectively. The DNN classifier also had the highest precision for Class 0 (94.07 percent) and an impressive F1-score for Class 1 (91.70 percent), indicating its ability to maintain a strong balance between precision and recall across both classes. Additionally, the DNN classifier exhibited remarkable computational efficiency, requiring only 0.09 million parameters and achieving an inference time of 3.2 milliseconds. This is significantly faster than YOLOv8-n and YOLOv8-l while using far fewer parameters, making it highly suitable for deployment on edge devices with limited computational resources. Results of the comparison are shown in Table 6.

Table 6 Condition Assessment Models Comparison: Methods

Method	Accuracy	P_0	P_1	R_0	R_1	F1_0	F1_1	Para (M)	Infer (ms)
Image input + YOLOv8-n	92.04	90.62	94.76	97.06	84.11	93.73	89.12	3.2	60.6
Image input + YOLOv8-l	92.79	91.67	94.87	97.06	86.05	94.29	90.24	43.7	124.4
Selected features + DNN classifier	93.56	94.07	91.34	94.50	92.06	94.74	91.70	0.09	3.2

6.3.2 Input Features Ablation Study

To further validate the importance of the selected features, an ablation study was conducted to analyze the impact of removing specific feature sets on model performance. **The results clearly showed that removing key features significantly degraded classification accuracy, reinforcing their importance in capturing the retroreflective and damage-related characteristics of traffic signs** (see Table 7).

Table 7 Input Features Ablation Study Results

Method	Accuracy	P_0	P_1	R_0	R_1	F1_0	F1_1	Para (M)	Infer (ms)
No GLCM	89.26	89.10	89.57	94.00	81.75	91.48	85.48	-	-
No GLCM, and Saturation	87.42	86.64	88.99	94.00	76.98	90.17	82.55	-	-
No GLCM, Saturation, Brightness, and Contrast	75.58	74.30	80.52	92.50	49.21	82.41	61.08	-	-
No GLCM, Saturation, Brightness, Contrast, and Hu moment	68.40	67.14	76.74	95.00	26.19	78.67	39.05	-	-

Removing GLCM features (texture-based descriptors) caused a notable drop in accuracy from 93.56 percent to 89.26percent, with the F1-score for Class 1 decreasing from 91.70 percent to 85.48 percent. This highlights the critical role of texture information in identifying damage, as worn-out or degraded signs often exhibit unique textural changes. Excluding GLCM and saturation features further reduced accuracy to 87.42 percent, with a decline in both precision and recall, particularly for damaged signs. This suggests that saturation contributes to the differentiation of faded or discolored signs, which are common indicators of deterioration. Removing GLCM, saturation, brightness, and contrast caused a significant drop in accuracy to 75.58 percent, with the F1-score for Class 1 dropping to 61.08 percent. This shows that brightness and contrast are essential for assessing sign visibility, as retroreflectivity deteriorates when signs become too dim or washed out. When all features except the category were removed, accuracy dropped to 68.40 percent, marking the lowest performance. The F1-score for Class 1 fell drastically to 39.05 percent, indicating that the model struggled to distinguish damaged signs from healthy ones without essential visual cues (see Figure 26).

The progressive decline in accuracy, precision, recall, and F1-score as more features were excluded demonstrated that a comprehensive feature set is necessary for achieving reliable traffic

sign condition assessment. Texture, saturation, brightness, and contrast collectively enhance the model's ability to detect signs of wear, fading, and retroreflectivity loss, all of which are critical factors in ensuring road safety.



Figure 26 Examples of Healthy and Damaged Traffic Signs

7.0 SYSTEM INTEGRATION AND SAMPLE DATA INVENTORY CREATION

7.1 OVERVIEW

This chapter focuses on the design and implementation of a comprehensive traffic sign management system, including detection, tracking, condition assessment, and inventory building. The first section provides an overview of the data flow through the system, followed by a detailed introduction to each module. The second section describes the sample data inventory visualization, which utilized the SRWeb Tool and highway datasets to construct an interactive inventory. The third section compares the traffic sign inventories generated by our system with those from WSDOT, analyzing key discrepancies and their implications. This structured approach ensured that the final system is both scalable and adaptable to real-world traffic sign management applications.

7.2 SYSTEM INTEGRATION

The system consists of several interconnected modules working together to achieve efficient traffic sign management. The process begins with input data from two primary sources: open street images and log video streams. These inputs are processed by the **Detection and Classification Module**, which outputs bounding box coordinates (Bbox) and classifications (Cls) for detected traffic signs. The **Tracking Module** then uses SORT tracking to follow signs across consecutive frames, ensuring continuity. Once the signs have been tracked, individual frames are passed to the **Assessment Module** for detailed evaluation. This module leverages advanced feature extraction methods, including GLCM features, Hu moments, brightness, contrast, and saturation, to assess the condition of each traffic sign. Simultaneously, GPS information is used to associate geospatial data with the processed signs. The outputs of the condition assessment module are combined into a unified traffic sign instance, including

classification, condition, and location data. These data are subsequently stored in an **Inventory Database**, which allows users to query and retrieve specific traffic sign information for further analysis or decision-making. Figure 27 shows the system architecture.

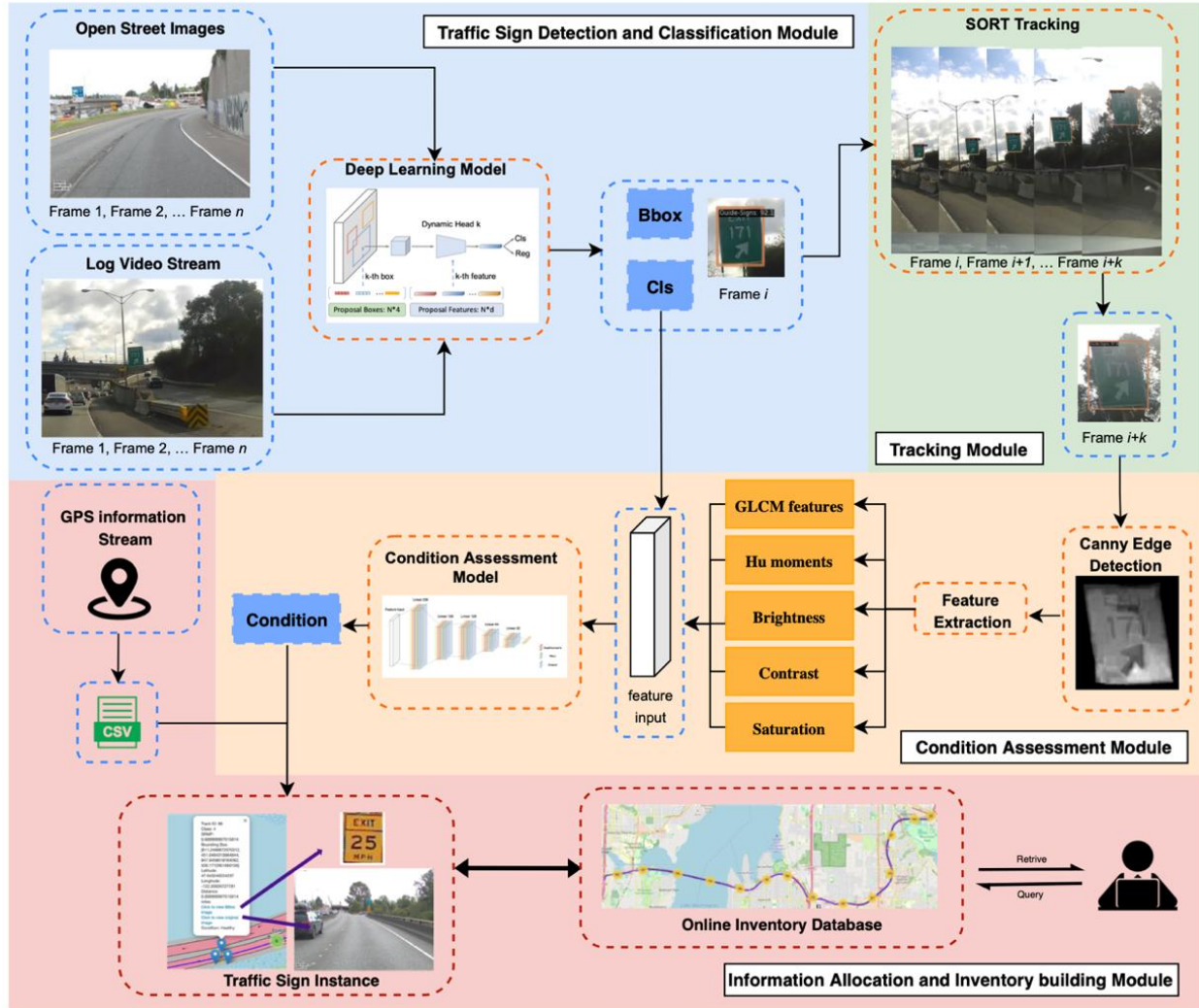


Figure 27 System Architecture

7.2.1 Detection and Classification Module

This module processes raw input data from open street images and log video streams, as depicted in the blue section of Figure 27. These inputs provide diverse traffic environments, ensuring that the model encounters varying lighting conditions, occlusions, and background clutter. A deep learning-based detection model, Sparse R-CNN, is employed for traffic sign

detection and classification. The model extracts bounding box coordinates (Bbox) and classification labels (Cls) for each detected traffic sign. Sparse R-CNN was selected for its robustness in handling challenging scenarios such as motion blur, partial occlusions, and signs positioned at varying distances.

A ResNet-101 backbone is utilized for feature extraction, leveraging its ability to capture rich hierarchical features to improve detection accuracy. The classification heads are customized to align with specific output requirements. The final output of this module includes the detected sign categories and their spatial coordinates (regions of interest, ROI) for every processed frame. The extracted ROIs are then forwarded to the tracking module for further processing to minimize redundancy in sign detection across frames.

7.2.2 Tracking Module

To efficiently process video inputs and ensure that each traffic sign is represented only once in our inventory, the SORT tracking algorithm is employed to minimize redundant outputs from the first module and to reduce the overall processing load on our system. In this module, every traffic sign is tracked until it leaves the video, and the final frame in which the sign can be tracked is utilized as the instance, mainly for two reasons. First, a larger ROI more clearly indicates any damage to the traffic sign, which can improve the performance of the subsequent assessment model. Second, the final detection of a traffic sign typically occurs when the data collection vehicle is about to pass the sign. At this point, the GPS data are most accurate, with minimal deviation from the actual coordinates of the sign, and can be considered a reliable estimate of its location. Consequently, the output from the tracking module is the cropped image of the last detected ROI and its corresponding frame number for each traffic sign.

7.2.3 Assessment Module

This module, represented in the orange section of Figure 27, evaluates the physical condition of traffic signs by analyzing extracted visual features. The assessment process starts with Canny Edge Detection, which isolates the sign from the background, reducing noise and improving feature extraction accuracy. Following this, several key visual features are extracted, including the following:

- **GLCM (Gray-Level Co-occurrence Matrix) Features:** Captures texture information to identify wear, fading, and degradation patterns.
- **Hu Moments:** Provides shape-based descriptors, which help in detecting physical deformations or misalignments.
- **Brightness:** Evaluates retroreflectivity performance, as faded or damaged signs often exhibit reduced brightness.
- **Contrast:** Measures the visibility of text and symbols against the background, ensuring the sign remains readable.
- **Saturation:** Assesses color intensity to detect fading, which is crucial for maintaining sign visibility.

A lightweight DNN classifier processes these extracted features, categorizing each sign as healthy or damaged. This model is designed for high accuracy with minimal computational overhead, making it suitable for real-time assessment on edge devices. The final condition classification result is then linked with the corresponding sign instance for inventory purposes.

7.2.4 Inventory Building Module

This module integrates condition assessment results, GPS data, and classification outputs to create a comprehensive traffic sign instance. Each instance includes the sign's type, condition,

and geographic location. The data are stored in an online inventory database, enabling easy retrieval and query functionality. Users can interact with the database to analyze traffic sign conditions, track historical changes, and optimize maintenance schedules.

This integrated system streamlines traffic sign management by combining advanced detection, tracking, and condition assessment techniques with efficient data storage and retrieval mechanisms. The modular design ensures scalability and adaptability, making it well-suited for deployment in diverse traffic environments.

7.3 SAMPLE DATA INVENTORY VISUALIZATION

To construct a comprehensive traffic sign inventory, we utilized the images from the SRweb Tool. The SRWeb Tool offers several advantages over self-collected data. It provides high-resolution imagery under optimal lighting conditions and ensures comprehensive route coverage, making it a valuable resource for constructing the inventory. Despite the tool's lower image refresh rate in comparison to self-collected video, the rich information attached to each image allowed us to effectively adapt our methodology. Then, we utilized another online dataset, ARM and SRMP Locations Along Washington Highway Dataset (Website: https://services.arcgis.com/IYrj3otxNjPsrTRD/arcgis/rest/services/WSDOT_State_Route_Points_Current/FeatureServer), to provide latitude and longitude information for every state route milepost (SRMP) along the Washington state highways (see Figure 28). By combining these two datasets, we were able to match the SRweb tool data with the corresponding geographic coordinates, enabling us to align the images with their precise locations.

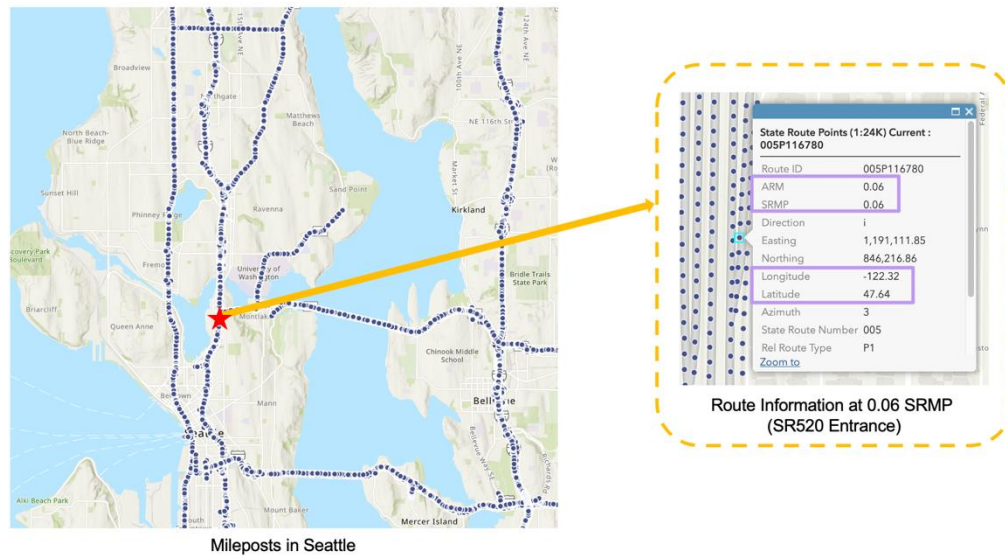


Figure 28 ARM and SRMP Locations Along Washington Highways

We used two specific routes—SR 520 and SR 530—as examples for building sample inventories. SR 520 is an urban road segment, covering 12.82 miles from I-5 to SR 202. This stretch comprises 1,283 images, and their locations are marked on a digital map with blue markers in Figure 29. Conversely, SR 530 is a rural road, specifically the segment from I-5 to Darrington, covering 32.89 miles. Over 3,000 images are included. This diversity ensured that our inventory-building process accounted for both urban and rural traffic conditions.

After the images have been processed, the generated traffic sign inventory is saved in a comma-separated values (CSV) format to ensure flexibility for transferring the data into other formats as needed. Additionally, the corresponding cropped traffic sign images are organized and saved in designated folders, as shown in Figure 30.

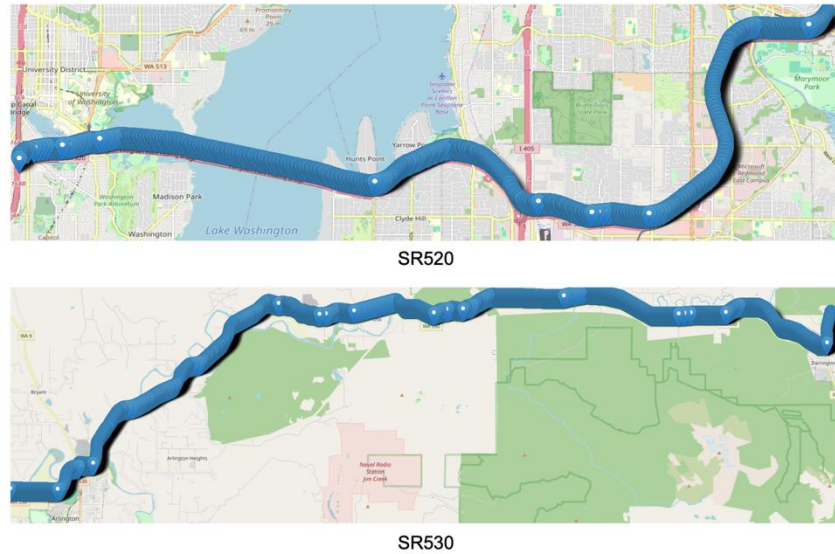


Figure 29 SR-520 and SR-530 Data and Its Locations

Image	Class Label	BBox (xmin, ymin, xmax, ymax)	Track ID	Image_int	OBJECTID	RouteID	ARM	SRMP	Direction	Easting	Northing
00000M.JPG	0	[170.10345458984375, 295.5492858896719, 261.61236572265625, 429.6379089355469]	1	0	1454906	520		0.0	0.0	1191140.69708739	846577.066148892
00000M.JPG	4	[475.8359006824219, 439.6993408203125, 521.1520385742188, 489.63653564453125]	2	0	1454906	520		0.0	0.0	1191140.69708739	846577.066148892
00000M.JPG	0	[169.39447021484375, 296.3263854980469, 257.2139892578125, 326.28717041015625]	3	0	1454906	520		0.0	0.0	1191140.69708739	846577.066148892
00001M.JPG	4	[143.83323669433594, 468.30224609375, 211.66134643554688, 541.8736572265625]	4	1	1455153	520	0.0099999997764825	0.0099999997764825	1	1191142.90180738	846631.07686764
00001M.JPG	4	[1184.607421875, 371.7438659667969, 1208.509765625, 403.9185485839844]	5	1	1455153	520	0.0099999997764825	0.0099999997764825	1	1191142.90180738	846631.07686764

 cropped_00000M_track_1_idx_0.jpg
 cropped_00000M_track_2_idx_1.jpg
 cropped_00001M_track_4_idx_3.jpg
 cropped_00001M_track_5_idx_4.jpg
 cropped_00001M_track_6_idx_5.jpg

Figure 30 Sample Data Inventory and Cropped Images

To enhance comprehension and accessibility, a visualized traffic sign inventory for SR 520 and SR 530 was created and is presented in a user-friendly website format and saved locally. Figures 31 and 32 show visualized maps in which detected traffic signs are displayed on the basis of their longitude and latitude. This spatial representation allows users to analyze the distribution of traffic signs along roadways, helping to identify high-density areas, missing signs, or locations requiring maintenance.

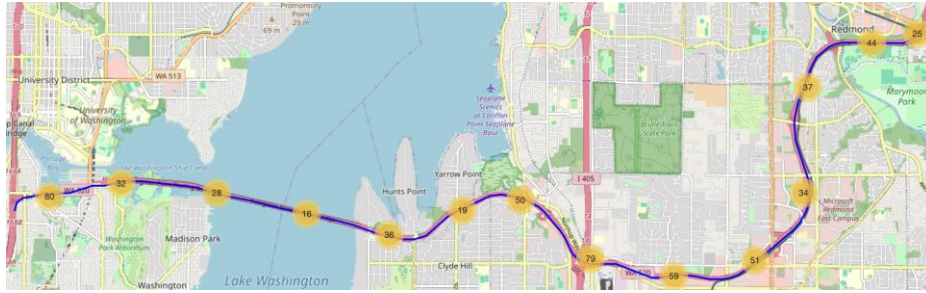


Figure 31 Inventory Map for SR 520

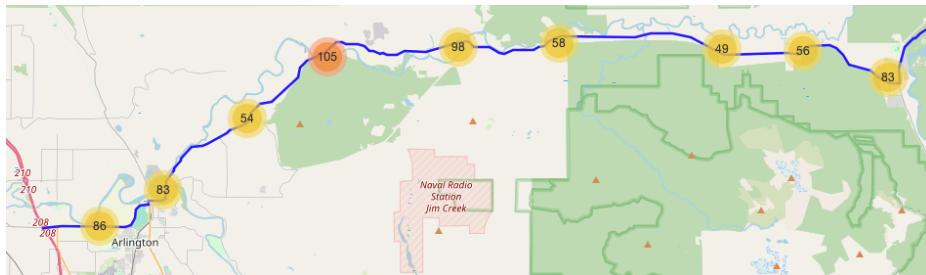


Figure 32 Inventory Map for SR 530

Figures 33 through 37 show the interactive interface developed for traffic sign monitoring and management. This interface integrates three key components: a filtering system, a visualized map, and an information panel. The filtering system enables users to refine traffic sign data on the basis of road name, traffic direction, and sign category, ensuring a focused and efficient analysis. The visualized map dynamically updates on the basis of selected filters, initially displaying an aggregate count of traffic signs along the route. As users zoom in, individual traffic signs appear as blue markers, allowing for detailed exploration. The information panel provides in-depth details about each traffic sign, including its tracking ID, classification, GPS coordinates, SRMP, cropped image, and condition assessment. By clicking on a marker, users can access a comprehensive breakdown of the selected sign's attributes. Each of the figures highlights different types of traffic signs. Figure 33 presents Guide and Information Signs. Figure 34 presents Warning Signs. Figure 35 displays Black and White Regulatory Signs. Figure 36 features Route Marker Signs. Figure 37 illustrates Marker Signs.

By integrating these elements into a user-friendly interface, the system provides an intuitive and efficient method for exploring traffic sign data. The ability to visually analyze sign placement, assess condition trends, and track maintenance needs makes this tool valuable for transportation agencies seeking to enhance roadway management and safety.

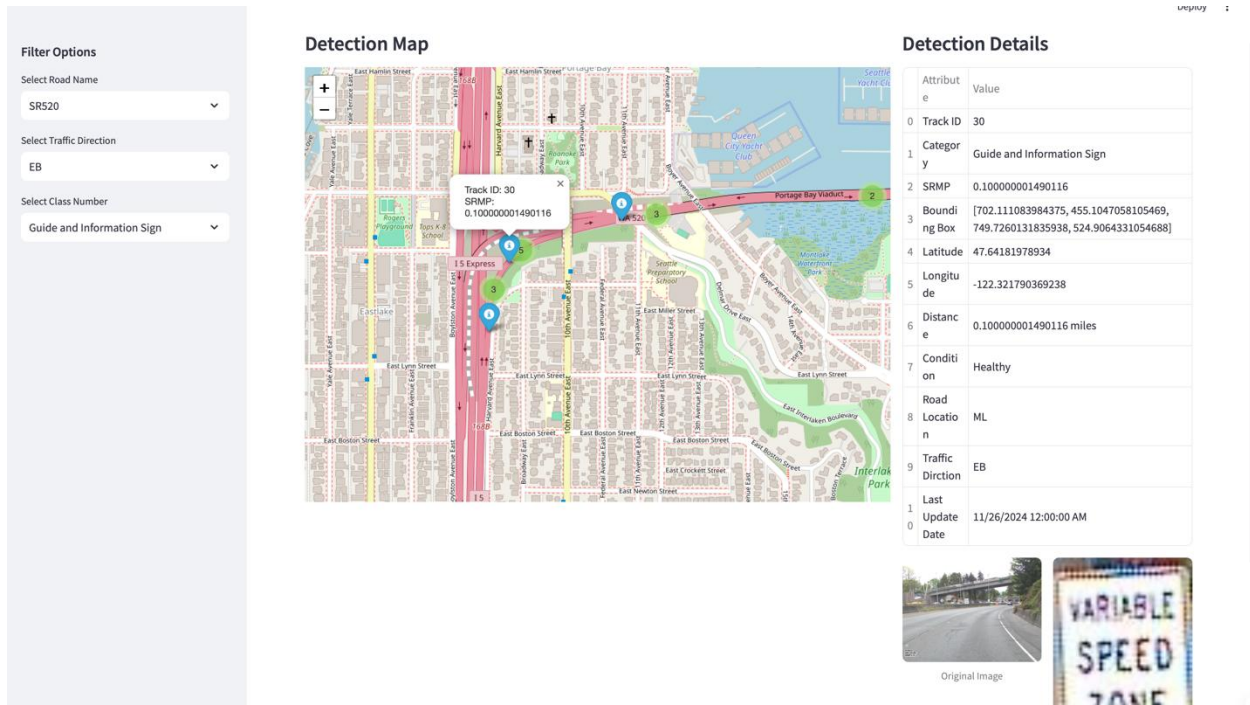


Figure 33 Visualized Inventory for SR 520 (Guide and Information Sign)

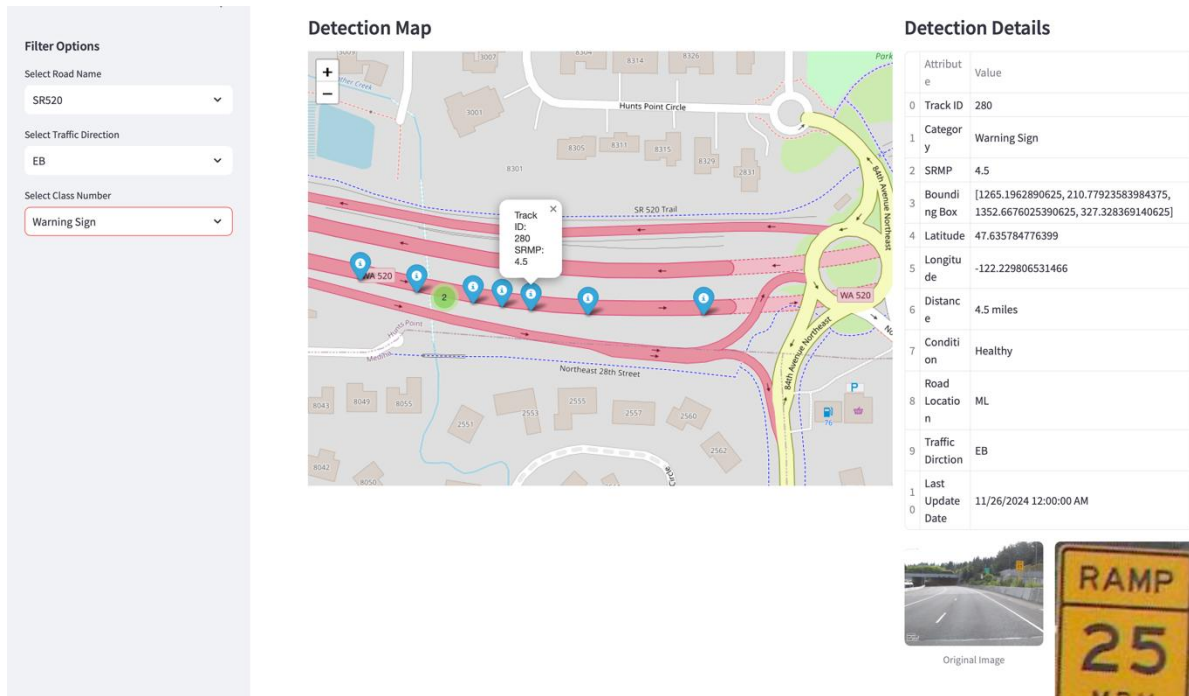


Figure 34 Visualized Inventory for SR 520 (Warning Sign)

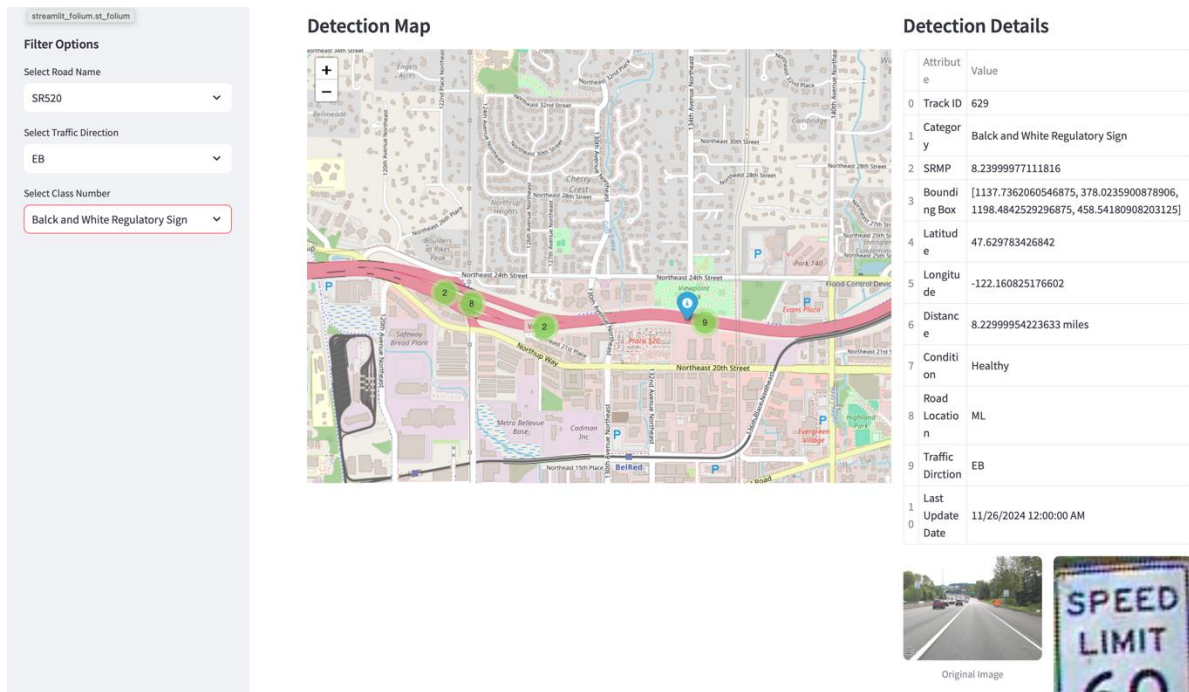


Figure 35 Visualized Inventory for SR 520 (Black and White Regulatory Sign)

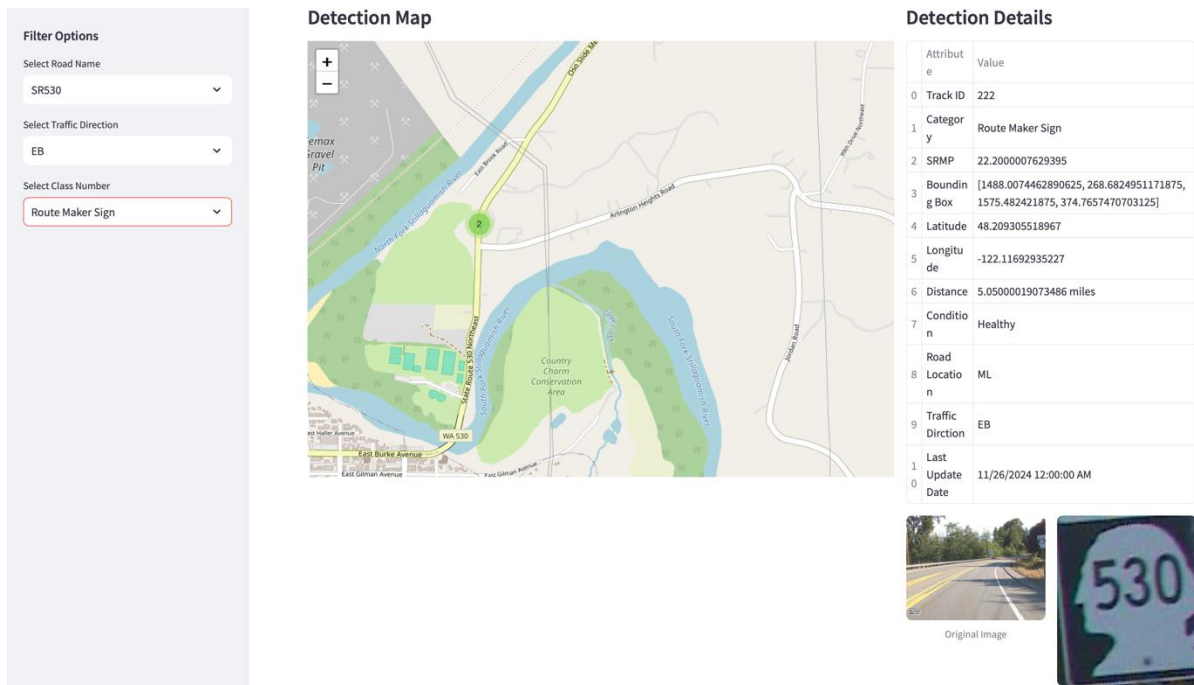


Figure 36 Visualized Inventory for SR 530 (Route Marker Sign)

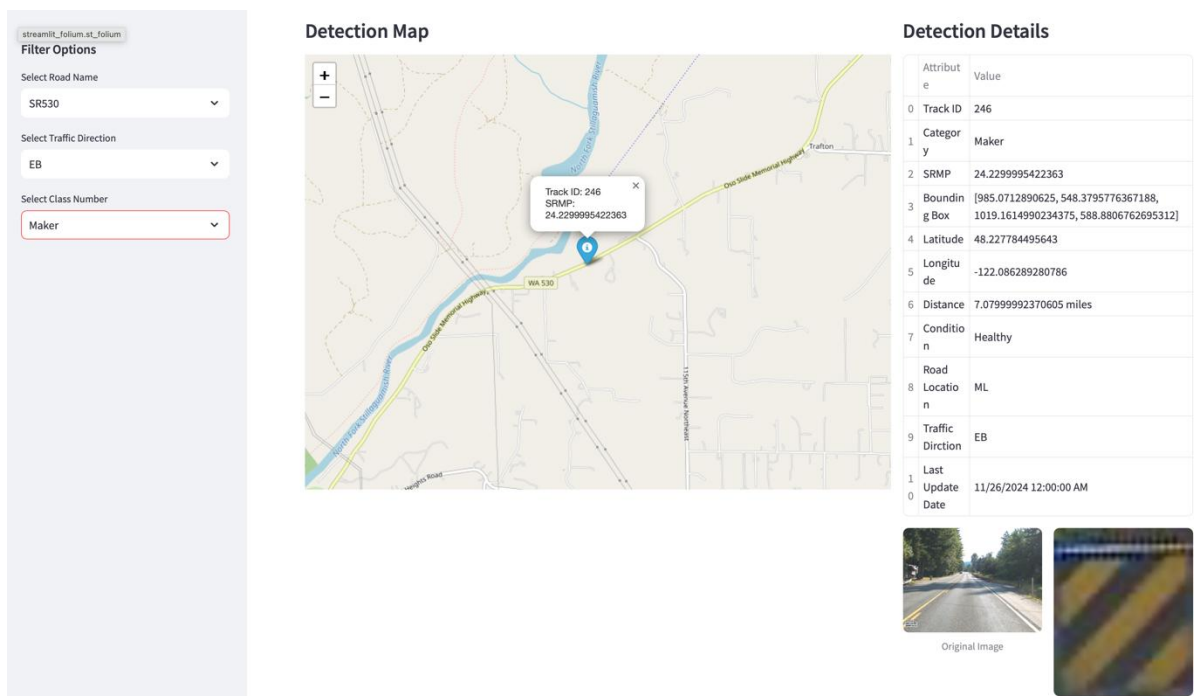


Figure 37 Visualized Inventory for SR 530 (Marker)

7.4 COMPARISON OF TRAFFIC SIGN INVENTORIES FOR SR 520 AND SR 530

This section presents the results of our traffic sign detection and classification system applied to the SR 520 and SR 530 highways and compares them with the WSDOT traffic sign inventory. Some of the detection results are shown in Figure 38.



Figure 38 Examples of Detection Results on SR 520

As shown in Figure 39, our system detected 145 out of 202 traffic signs on the SR 520 eastbound main lane, achieving a detection rate of 71.78 percent, and 440 out of 602 traffic signs on SR 530 eastbound main lane, resulting in a 73.09 percent detection rate. These rates indicated reasonable performance but also highlighted notable **discrepancies** that require further analysis. Several factors contributed to the differences between our results and the WSDOT inventory.

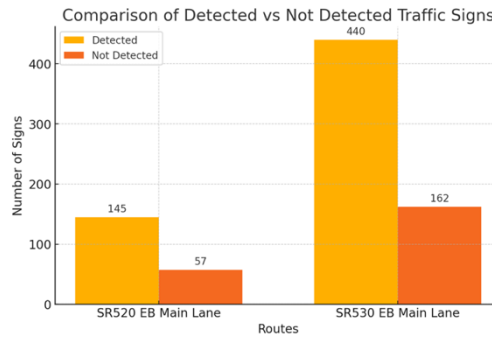


Figure 39 Comparison of Detected vs Not Detected Traffic Signs

First, milepost labeling rules differ between our system and WSDOT. Our system assigns the milepost on the basis of the last frame where the traffic sign was detected, whereas WSDOT follows a different convention. As a result, traffic signs in our system may appear at nearby mileposts rather than the exact locations specified in WSDOT’s records, complicating the automated matching process. Figure 40 illustrates a traffic sign appearing in both WSDOT’s official documentation and our developed inventory. The upper table represents the WSDOT traffic sign record, while the lower interface displays our results. In comparison to WSDOT’s table-based approach, our system provides more direct visualization with precise GPS coordinates, enhancing localization accuracy and sign management.

Sign ID	Region	Maintenance Area	State Route	Mile Post	Special Mile Post	Road Location	Traffic Direction	Prog Code	Sort Order	Fabrication Number	Sign Message	Last Update Date	Install Date
236713	1	5	520	0.030	0.000	ML	EB	0	1	I6-901	ADOPT-A-HIGHWAY	06/13/2013 2:33:21 PM	03/17/2013 12:00:00 AM

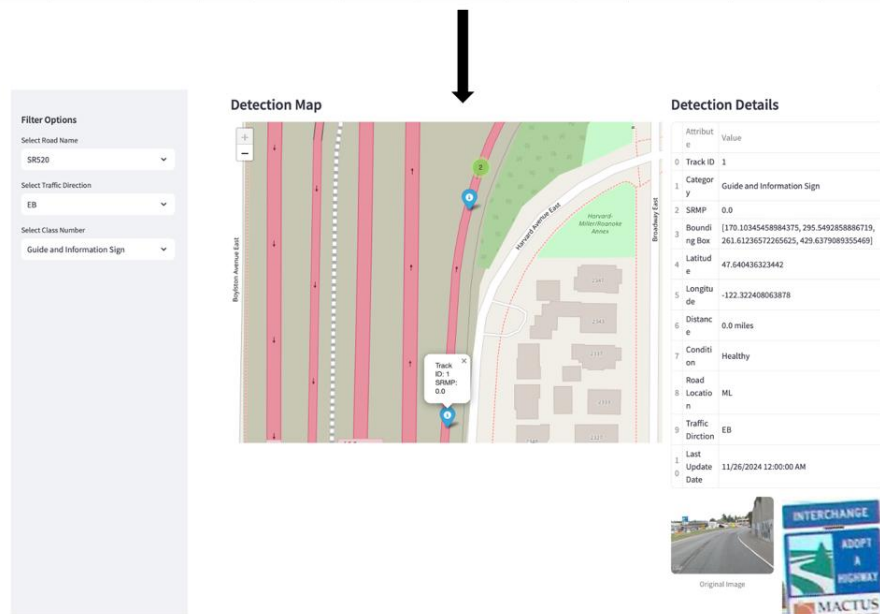


Figure 40 A Traffic Sign Instance Both Appearing in WSDOT (Top) and Our Inventories (Bottom)

Second, some traffic signs in the WSDOT dataset fall outside the SRweb image scope, particularly those located on ramps, overpasses, and intersections. Because our system relies on SRweb images for detection, these signs are not included in our results, leading to occasional gaps. However, if additional data sources, such as dash camera log videos or external image datasets, are integrated into the system, then these missing traffic signs can be detected, and the inventory can be updated accordingly. This flexibility allows for continuous improvements in sign coverage and inventory completeness.

Lastly, older traffic signs are more likely to be missing from our detection results because of replacements or removals over time. For example, as shown in Figure 41, some traditional speed limit signs have been replaced with digital signs, yet these updates are not reflected in the inventory. Figure 42 reveals that 53 percent of detected signs were installed between 1980 and 2000, whereas 68 percent of undetected signs dated back to the same period,

supporting the notion that older signs are more frequently missing. This finding underscores the importance of regularly updating the traffic sign inventory to ensure that it accurately reflects the current state of roadway signage.

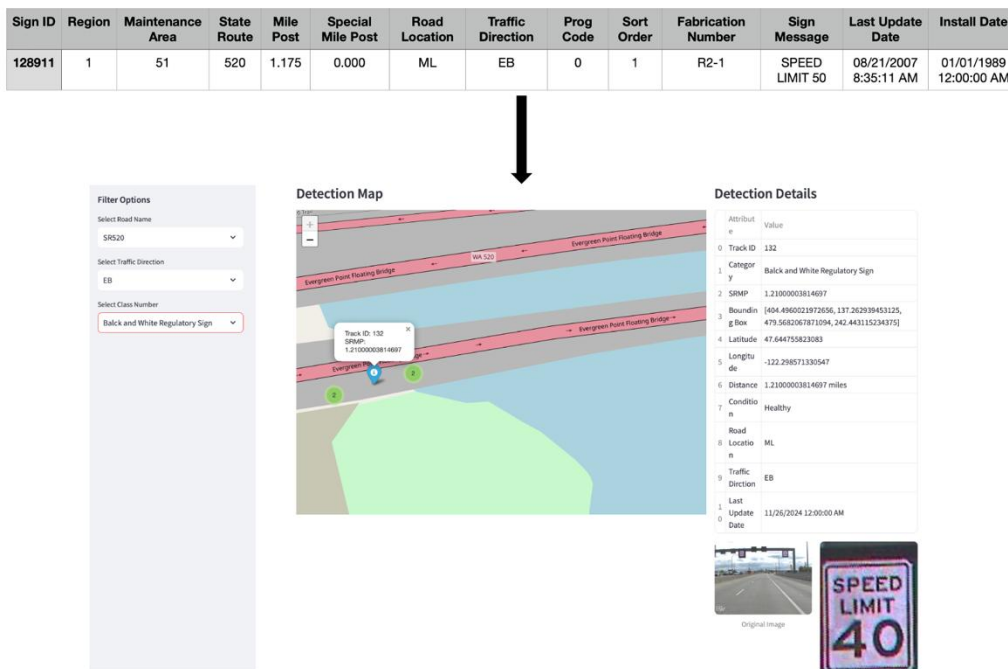


Figure 41 A Digital Speed Limit Sign Only Showing in Our Inventory (Bottom)

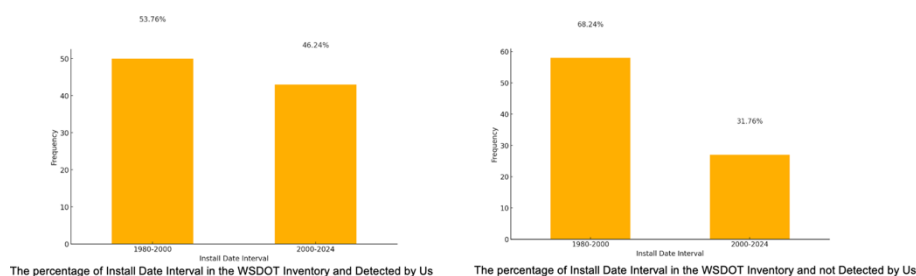


Figure 42 Traffic Sign Installation Date Comparison

As shown by the comparison, our system offers several key advantages that can support and enhance the official traffic sign inventory.

One advantage is that our system provides a continuously updated inventory, ensuring that recent roadway changes are accurately reflected. Unlike manually built inventories, which

may not always capture modifications in signage, our approach enables real-time and efficient detection of updates. By maintaining a regularly updated database, our system ensures that roadway signage data remain accurate and reliable, helping transportation agencies keep pace with evolving infrastructure.

Another advantage is precise geographic information. Our system enhances location accuracy by providing exact latitude and longitude coordinates for each detected traffic sign, significantly improving upon the previous milepost-based labeling approach. Milepost-based records can introduce displacement errors, particularly near milepost transitions, interchanges, and overpasses, where precise location tracking is crucial. By integrating highly accurate spatial data, our system reduces confusion and enhances the usability of the traffic sign inventory for mapping, navigation, and roadway management.

Additionally, our system offers comprehensive condition assessments, providing critical insights that go beyond simple traffic sign detection and classification. By incorporating condition evaluations into the inventory, our system enables proactive maintenance planning, allowing transportation agencies to prioritize sign replacements and repairs. This can help ensure that all roadway signs remain effective, visible, and compliant with safety standards, ultimately improving overall traffic management and road safety.

In conclusion, the discrepancies between our detected traffic signs and the WSDOT inventory emphasize the critical need for regularly updating the official inventory to ensure accuracy and reliability. As traffic signs are continually replaced, removed, or newly installed, maintaining an up-to-date database is essential for effective roadway management and safety monitoring. Our findings demonstrate that an automated system like ours can complement manual inventories by providing real-time updates, precise geolocation data, and comprehensive

condition assessments. By integrating such advancements, transportation agencies can enhance the accuracy of their traffic sign records, optimize maintenance efforts, and improve overall roadway efficiency and safety.

8.0 CONCLUSIONS AND FUTURE WORK

8.1 CONCLUSION

This research developed a robust and automated approach to traffic sign detection, recognition, and condition assessment by leveraging computer vision and deep learning methodologies. The study demonstrated the effectiveness of learning-based models in automating traffic sign inventory management, which offer significant advantages over traditional manual and semi-automated approaches. By integrating public datasets with self-collected and open-source imagery, the developed system ensures adaptability to diverse environmental conditions and sign variations, enhancing its applicability in real-world scenarios.

In Sections 1 and 2 we introduced the motivation for this research, emphasizing the need for automated traffic sign management because of the limitations of manual assessments, such as inefficiency, inconsistency, and high labor costs. We also provided a comprehensive literature review, covering traditional and modern deep learning-based traffic sign detection methods, traffic sign inventory building systems, and automated traffic sign condition assessment methods.

Section 3 focused on dataset selection and verification, detailing the process of integrating public datasets, open street imagery, and self-collected data to construct a comprehensive and diverse dataset. By ensuring data quality and coverage, we enhanced the robustness of the detection and classification models, enabling them to handle real-world complexities such as varying lighting conditions, occlusions, and environmental noise.

In Section 4 we discussed the traffic sign data collection and dataset construction process. A data collection system was deployed using an edge-based vanilla traffic sign detection model, ensuring that data were efficiently captured from roadway environments. This step was critical in

bridging the gap between real-world deployment and existing datasets, which often lack sufficient variation in road conditions and sign placements.

In Section 5 we outlined the development and evaluation of a learning-based traffic sign detection and recognition algorithm. After assessing multiple state-of-the-art models, we selected Sparse R-CNN for its optimal balance between detection accuracy and computational efficiency. The model demonstrated robust performance in complex environments, accurately identifying traffic signs even under challenging conditions such as occlusion, glare, and poor visibility. Beyond detection, in Section 6 we also focused on traffic sign condition assessment, utilizing image-based methods and deep learning techniques to evaluate retroreflectivity and maintenance status. By incorporating key visual features such as GLCM texture analysis, brightness, contrast, and saturation, we developed an effective classification model to determine whether traffic signs require maintenance. The condition assessment system provides critical insights that enable transportation agencies to prioritize sign maintenance efforts, ultimately enhancing road safety and regulatory compliance.

In Section 7 we integrated the developed system with sample data inventory creation, demonstrating its practical applicability for large-scale deployment. By incorporating GPS data, classification results, and condition assessment outputs, we built an interactive traffic sign inventory that allows for real-time monitoring, query-based retrieval, and historical tracking of traffic signs. The system supports dynamic updates, ensuring that newly collected data can continuously refine and expand the inventory.

Overall, this research addressed key challenges in traffic sign inventory management, providing a scalable, cost-effective, and automated solution for transportation agencies. By eliminating the need for manual inspections, the proposed system will significantly reduce

human labor costs, increase inspection efficiency, and enhance data accuracy. Furthermore, by automating detection, classification, and condition assessment, this research contributes to the advancement of ITS and promotes the broader adoption of CV in roadway infrastructure management. The modular design of the system ensures that it can be easily adapted to different geographic regions, integrated with new data sources, and scaled for future transportation management needs.

8.2 FUTURE WORK

Despite the promising results achieved in this study, several areas remain open for future improvement and exploration:

- **Enhancing Small Object Detection** – While Sparse R-CNN performed well, small traffic signs and partially occluded signs remain challenging. Future research could explore transformer-based models or multi-scale feature extraction techniques to improve detection in complex environments.
- **Real-Time Deployment Optimization** – The current system operates efficiently, but further optimization, such as model quantization or edge computing integration, could improve real-time performance on low-power devices.
- **Expansion of Condition Assessment Capabilities** – Incorporating additional sensor data, such as LiDAR or infrared imaging, could enhance the accuracy of traffic sign reflectivity and degradation assessments, leading to more comprehensive maintenance strategies.
- **Integration with Geographic Information Systems (GIS)** – Future work could integrate the traffic sign detection and recognition system with GIS platforms for real-

time mapping and visualization, allowing for better spatial analysis and decision-making for transportation agencies.

- **Automated Labeling for Large-Scale Datasets** – While the current dataset includes diverse sources, semi-supervised or self-supervised learning techniques could be explored to reduce manual annotation efforts and improve dataset scalability.
- **Long-Term Monitoring and Predictive Maintenance** – Implementing a temporal analysis framework that tracks changes in traffic sign conditions over time would allow for predictive maintenance strategies, enabling agencies to schedule repairs before signs reached critical degradation levels.

By continuing to refine and expand this approach, transportation agencies can achieve safer, more efficient roadway networks, reinforcing the importance of data-driven infrastructure management in modern urban and highway planning.

9.0 REFERENCES

- Acilo, J.P.N., A.G.S.D. Cruz, M.K.L. Kaw, M.D. Mabanta, V.G.G. Pineda, and E.A. Roxas, "Traffic sign integrity analysis using deep learning," in 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA). IEEE, 2018, pp. 107–112.
- ADOT, "Roadway-Maintenance-Summary-Report-October-2019.pdf." Accessed: Feb. 06, 2024. [Online]. Available: <https://azmag.gov/Portals/0/Documents-Ext/FLCP/Roadway-Maintenance-Summary-Report-October-2019.pdf>
- Ai, C., and Y. J. Tsai, "An automated sign retroreflectivity condition evaluation methodology using mobile LIDAR and computer vision," *Transportation Research Part C: Emerging Technologies*, vol. 63, pp. 96–113, Feb. 2016, doi: 10.1016/j.trc.2015.12.002.
- Ai, C., and Y. Tsai, "Critical Assessment of an Enhanced Traffic Sign Detection Method Using Mobile LiDAR and INS Technologies," *Journal of Transportation Engineering*, Dec. 2014, doi: 10.1061/(ASCE)TE.1943-5436.0000760.
- Almutairy, F., Alshaabi, T., Nelson, J., & Wshah, S., "Arts: Automotive repository of traffic signs for the United States" in *IEEE Transactions on Intelligent Transportation Systems*, 2019, 22(1), 457-465.
- Alnasser, A., Sun, H., and Jiang, J. (2019). "Cyber Security Challenges and Solutions for V2X Communications: A Survey." arXiv:1901.01053, Submitted to *Computer Networks*.
- Møgelmoose, Andreas, Dongran Liu, and Mohan Manubhai Trivedi. "Detection of US traffic signs." *IEEE Transactions on Intelligent Transportation Systems* 16.6 (2015): 3116-3125.
- ASCE, "ASCE's 2021 American Infrastructure Report Card | GPA: C-," ASCE's 2021 Infrastructure Report Card |. Accessed: Feb. 06, 2024. [Online]. Available: <https://infrastructurereportcard.org/>
- Balali, V., Ashouri Rad, A. & Golparvar-Fard, M. Detection, classification, and mapping of U.S. traffic signs using google street view images for roadway inventory management. *Vis. in Eng.* 3, 15, 2015
- Balali, V., M.A. Sadeghi, and M. Golparvar-Fard, "Image-based retro-reflectivity measurement of traffic signs in day time," *Advanced Engineering Informatics*, vol. 29, no. 4, pp. 1028–1040, Oct. 2015, doi: 10.1016/j.aei.2015.08.003.
- Bangquan, X.; Xiong, W.X. Real-time embedded traffic sign recognition using efficient convolutional neural network. *IEEE Access* 2019, 7, 53330–53346.
- Barnes, N. and A. Zelinsky, "Real-time radial symmetry for speed sign detection", *IEEE Intelligent Vehicles Symposium* 2004, pp. 566-571, 2004.
- Bochkovskiy A, Wang C Y, Liao H Y M. "Yolov4: Optimal speed and accuracy of object detection". arXiv preprint arXiv:2004.10934, 2020.

- Boggs, W., Heaslip, K., & Louisell, C.. Analysis of sign damage and failure: Utah case study. *Transportation research record*, 2013, 2337(1), 83-89.
- Cai W, Li J, Xie Z, et al. "Street object detection based on faster R-CNN", in 2018 37th Chinese Control Conference (CCC). IEEE, 2018: 9500-9503.
- Chen, T., and J. Ren, "Mfl-yolo: An object detection model for damaged traffic signs," arXiv preprint arXiv:2309.06750, 2023.
- CTDOT "2022 Highway Transportation Asset Management Plan" Accessed: Feb. 06, 2024. [Online]. Available: <https://portal.ct.gov/dot/-/media/dot/tam/transportation-asset-management-plan-fhwa-certified-9302022.pdf?rev=989262de6f41476f8365f271f293ec6f&hash=21AC54AF2CCA5CE106855062270548E9>
- Dai J, Li Y, He K, et al. "R-FCN: Object detection via region-based fully convolutional networks", in *Advances in neural information processing systems*, 2016, 29.
- Ertler, C., J. Mislej, T. Ollmann, L. Porzi, G. Neuhold, and Y. Kuang, "The mapillary traffic sign dataset for detection and classification on a global scale," in *European Conference on Computer Vision*. Springer, 2020, pp. 68–84.
- Fatmehsari, Y.R., A. Ghahari and R. A. Zoroofi, "Gabor wavelet for road sign detection and recognition using hybrid classifier", 2010 *International Conference on Multimedia Computing and Information Technology (MCIT)*, pp. 25-28, 2010.
- Fernando, W.H.D., and S. Sotheeswaran, "Automatic road traffic signs detection and recognition using 'You Only Look Once' version 4 (YOLOv4)," 2021 *International Research Conference on Smart Computing and Systems Engineering (SCSE)*, Colombo, Sri Lanka, 2021, pp. 38-43
- Garvila, D.M., "Traffic sign recognition revisited", *Informatik aktuell Mustererkennung* 1999, pp. 86-93, 1999.
- Ge, Z., S. Liu, F. Wang, Z. Li, and J. Sun, "Yolox: Exceeding yolo series in 2021". arXiv preprint arXiv:2107.08430, 2021.
- González, Á., Garrido, M. Á., Llorca, D. F., Gavilán, M., Fernández, J. P., Alcantarilla, P. F., ... & de Toro, P. R.. Automatic traffic signs and panels inspection system using computer vision. *IEEE Transactions on intelligent transportation systems*, 2011, 12(2), 485-499.
- Gracia-Garrido, M., M. Sotelo and E. Martin-Gorostiza, "Fast traffic sign detection and recognition under challenging lighting conditions", 2006 *IEEE Intelligent Transportation Systems Conference* 811–816, 2006.
- Haralick, R. M., K. Shanmugam, and I. H. Dinstein, Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, , No. 6, 1973, pp. 610–621

- Houben, S., J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The German traffic sign detection benchmark," in The 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX, USA: IEEE, Aug. 2013, pp. 1–8. doi: 10.1109/IJCNN.2013.6706807.
- Hu, M.-K., Visual pattern recognition by moment invariants. IRE transactions on information theory, Vol. 8, No. 2, 1962, pp. 179–187
- Khalilikhah, M., and K. Heaslip, "Analysis of factors temporarily impacting traffic sign readability," International Journal of Transportation Science and Technology, vol. 5, no. 2, pp. 60–67, Oct. 2016, doi: 10.1016/j.ijtst.2016.09.003.
- Khalilikhah, M., K. Heaslip, and K. Hancock, "Traffic sign vandalism and demographics of local population: A case study in Utah," Journal of Traffic and Transportation Engineering (English Edition), vol. 3, no. 3, pp. 192–202, Jun. 2016, doi: 10.1016/j.jtte.2015.11.001.
- KOMO News, T. M. Reporter, "Graffiti covering freeway and road signs in Seattle draws concern from drivers," KOMO. Accessed: Feb. 06, 2024. [Online]. Available: <https://komonews.com/news/local/graffiti-covering-freeway-road-signs-seattle-draws-concern-drivers-traffic-downtown-mayor-bruce-harrel-wsdot-washington-transportation-king-county-vandalism-signs-tagging-driving-road-safety>
- Li, J.; Wang, Z. Real-time traffic sign recognition based on efficient CNNs in the wild. IEEE Trans. Intell. Transp. Syst. 2018.
- Liu W, Anguelov D, Erhan D, et al. "SSD: Single shot multibox detector" in Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14.
- Magnussen, A.F., N. Le, L. Hu, and W. Eric Wong, "A Survey of the Inadequacies in Traffic Sign Recognition Systems for Autonomous Vehicles," Int J Performability Eng, vol. 16, no. 10, p. 1588, 2020, doi: 10.23940/ijpe.20.10.p10.15881597.
- Møgelmoose, A., D. Liu and M.M. Trivedi, "Detection of U.S. Traffic Signs," in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 6, pp. 3116–3125, 2015
- Nguwi, Y., and A. Kouzani, "Automatic road sign recognition using neural networks", The 2006 IEEE International Joint Conference on Neural Network Proceedings, pp. 3955–3962, 2006.
- Oruklu, E., D. Pesty, J. Neveux and J. Guebey, "Real-time traffic sign detection and recognition for in-car driver assistance systems", 2012 IEEE 55th International Midwest Symposium on Circuits and Systema (MWSCAS), pp. 976–979, 2012.
- Philipsen, M.P., M.B. Jensen, A. Mogelmoose, T.B. Moeslund, and M.M. Trivedi, "Traffic Light Detection: A Learning Algorithm and Evaluations on Challenging Dataset," in 2015 IEEE 18th International Conference on Intelligent Transportation Systems, Gran Canaria, Spain: IEEE, Sep. 2015, pp. 2341–2345. doi: 10.1109/ITSC.2015.378.

- Qin, F., B. Fang and H. Zhao, "Traffic sign segmentation and recognition in scene images", 2010 Chinese Conference on Pattern Recognition (CCPR), pp. 1-5, 2010.
- Radoš, K., J. Downes, D.-S. Pham, and A. Krishna, "End-to-End Traffic Sign Damage Assessment," in 2022 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Sydney, Australia: IEEE, Nov. 2022, pp. 1–8. doi: 10.1109/DICTA56598.2022.10034587.
- Ré, J.M., J. D. Miles, and P. J. Carlson, "Analysis of In-Service Traffic Sign Retroreflectivity and Deterioration Rates in Texas," *Transportation Research Record*, vol. 2258, no. 1, pp. 88–94, Jan. 2011, doi: 10.3141/2258-11.
- Ren, S, He K, Girshick R, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks", in *Advances in neural information processing systems*, 2015, 28.
- Romdhane, N.B., H. Mliki and M. Hammami, "An improved traffic signs recognition and tracking method for driver assistance system", 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS) 1–6, 2016.
- Shustanov, A.; Yakimov, P. CNN design for real-time traffic sign recognition. *Procedia Eng.* 2017, 201, 718–725.
- Siegmann, P., R. J. Lopez-Sastre, P. Gil-Jimenez, S. Lafuente-Arroyo, and S. Maldonado-Bascon, "Fundamentals in Luminance and Retroreflectivity Measurements of Vertical Traffic Signs Using a Color Digital Camera," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 3, pp. 607–615, Mar. 2008, doi: 10.1109/TIM.2007.911643.
- Sun, Y., P. Ge and D. Liu, "Traffic Sign Detection and Recognition Based on Convolutional Neural Network," 2019 Chinese Automation Congress (CAC), Hangzhou, China, 2019, pp. 2851-2854
- Sun, P., R. Zhang, Y. Jiang, T. Kong, C. Xu, W. Zhan, M. Tomizuka, L. Li, Z. Yuan, C. Wang, et al., Sparse r-cnn: End-to-end object detection with learnable proposals. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 14454–14463.
- Supreeth, H.s.g., and C. M. Patil, "An approach towards efficient detection and recognition of traffic signs in videos using neural networks", 2016 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), 2016.
- Temel, D., T. Alshaw, M.-H. Chen, and G. AlRegib, "Challenging Environments for Traffic Sign Detection: Reliability Assessment under Inclement Conditions." *arXiv*, Aug. 28, 2019. Accessed: Nov. 27, 2023. [Online]. Available: <http://arxiv.org/abs/1902.06857>
- Xing, M., M. Chunyang, W. Yan, W. Xiaolong and C. Xuetao, "Traffic sign detection and recognition using color standardization and Zernike moments", 2016 Chinese Control and Decision Conference (CCDC), pp. 5195-5198, 2016.

- You, C., C. Wen, H. Luo, C. Wang, and J. Li, “Rapid traffic sign damage inspection in natural scenes using mobile laser scanning data,” in 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX: IEEE, Jul. 2017, pp. 6271–6274. doi: 10.1109/IGARSS.2017.8128440.
- Zaibi, A.; Ladgham, A.; Sakly, A. A lightweight model for traffic sign classification based on enhanced LeNet-5 network. *J. Sensors* 2021, 2021, 8870529
- Zhang, J.; Wang, W.; Lu, C.; Wang, J.; Sangaiah, A.K. Lightweight deep network for traffic sign classification. *Ann. Telecommun.* 2020, 75, 369–379.
- Zheng, Z., H. Zhang, B. Wang and Z. Gao, "Robust traffic sign recognition and tracking for Advanced Driver Assistance Systems", 2012 15th International IEEE Conference on Intelligent Transportation Systems, pp. 704-709, 2012.
- Zhu, X., W. Su, L. Lu, B. Li, X. Wang, and J. Dai, Deformable DETR: Deformable Transformers for End-to-End Object Detection. In *International Conference on Learning Representations*, 2021.
- Zhu, Y.; Yan, W.Q. Traffic sign recognition based on deep learning. *Multimed. Tools Appl.* 2022, 81, 17779–17791.
- Zhu, Z., D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, “Traffic-Sign Detection and Classification in the Wild,” in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 2110–2118. doi: 10.1109/CVPR.2016.232.

Title VI Notice to Public

It is the Washington State Department of Transportation's (WSDOT's) policy to assure that no person shall, on the grounds of race, color, or national origin, as provided by Title VI of the Civil Rights Act of 1964, be excluded from participation in, be denied the benefits of, or be otherwise discriminated against under any of its programs and activities. Any person who believes his/her Title VI protection has been violated, may file a complaint with WSDOT's Office of Equity and Civil Rights (OECR). For additional information regarding Title VI complaint procedures and/or information regarding our non-discrimination obligations, please contact OECR's Title VI Coordinator at (360) 705-7090.

Americans with Disabilities Act (ADA) Information

This material can be made available in an alternate format by emailing the Office of Equity and Civil Rights at wsdotada@wsdot.wa.gov or by calling toll free, 855-362-4ADA(4232). Persons who are deaf or hard of hearing may make a request by calling the Washington State Relay at 711.