CPBS

Center for Pedestrian and Bicyclist Safety

A Tier-1 University Transportation Center

Integrating Non-Motorist Facility Data into Comprehensive Road Safety Assessment

July 2024	A Report From the Center for Pedestrian and Bicyclist Safety

Xiao Qin
University of Wisconsin-Milwaukee (UWM)

Robert Schneider
University of Wisconsin-Milwaukee (UWM)

Md Abu Sayed
University of Wisconsin-Milwaukee (UWM)

Mohammad Abrari Vajari
University of Wisconsin-Milwaukee (UWM)

About the Center for Pedestrian and Bicyclist Safety (CPBS)

The Center for Pedestrian and Bicyclist Safety (CPBS) is a consortium of universities committed to eliminating pedestrian and bicyclist fatalities and injuries through cutting-edge research, workforce development, technology transfer, and education. Consortium members include: The University of New Mexico; San Diego State University; The University of California Berkeley; The University of Tennessee Knoxville; and The University of Wisconsin Milwaukee. More information can be found at: https://pedbikesafety.org

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

The U.S. Department of Transportation requires that all University Transportation Center reports be published publicly. To fulfill this requirement, the Center for Pedestrian and Bicyclist Safety provides reports publicly on its website, www.pedbikesafety.org The authors may copyright any books, publications, or other copyrightable materials developed in the course of, or under, or as a result of the funding grant; however, the U.S. Department of Transportation reserves a royalty-free, nonexclusive and irrevocable license to reproduce, publish, or otherwise use and to authorize others to use the work for government purposes.

Acknowledgments

This study was funded, partially or entirely, by a grant from the Center for Pedestrian and Bicyclist Safety (CPBS), supported by the U.S. Department of Transportation (USDOT) through the University Transportation Centers program. The authors would like to thank CPBS and the USDOT for their support of university-based research in transportation, and especially for the funding provided in support of this project.

TECHNICAL DOCUMENTATION

1. Project No.	2. Government Accession No.	3. Recipient's Catalog No.
23UWM02		

4. Title and Subtitle	5. Report Date
Integrating Non-Motorist Facility Data into Comprehensive Road Safety Assessment	July 2024
	6. Performing Organization Code
	N/A
7. Author(s)	**8. Performing Organization Report No.**
Xiao Qin https://orcid.org/0000-0002-2299-1656 Robert Schneider https://orcid.org/0000-0002-6225-3615 Md Abu Sayed https://orcid.org/0000-0003-0106-3081 Mohammad Abrari Vajari https://orcid.org/0000-0001-6150-1077	N/A
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
Center for Pedestrian and Bicyclist Safety Centennial Engineering Center 3020 The University of New Mexico Albuquerque, NM 87131	**11. Contract or Grant No.** 69A3552348336
12. Sponsoring Agency Name and Address	**13. Type of Report and Period Covered**
United States of America Department of Transportation Office of Research, Development, and Technology (RD&T)	Final Report – June 2023 to May 2024 **14. Sponsoring Agency Code** USDOT OST-R

15. Supplementary Notes

Report accessible via the CPBS website https://pedbikesafety.org and DOI https://doi.org/10.21949/sjgp-f537

16. Abstract

As active transportation gains popularity for both work and leisure trips, the need for reliable and accurate data on non-motorist activities and infrastructure becomes increasingly critical. Traditionally, detailed data on non-motorist activities and facilities has been scarce and challenging to obtain. However, advancements in technology now provide an opportunity to collect these data. This study aims to investigate recent technological advancements and methodologies for collecting non-motorist data and evaluate their ability to detect, classify, monitor changes in, and assess the condition of these facilities. A key focus is the use of vision-based large language models, specifically ChatGPT, for retrieving and evaluating non-motorist facilities from aerial images. This study tested ChatGPT's performance on a dataset containing satellite images to evaluate its ability to accurately detect and identify features such as pedestrian crosswalks. The results showed that ChatGPT can reliably assess crosswalk and sidewalk conditions, informing maintenance and improvement strategies. However, limitations were observed, including inconsistent detection of bicycle lanes without explicit visual cues and challenges in classifying crosswalks into specific types. Additionally, it struggles with reliably identifying street lighting and pedestrian signals, which are critical for comprehensive safety assessments.

17. Key Words	18. Distribution Statement
Non-motorized transportation; Aerial photography; Road markings; Artificial intelligence; Equity	No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	45	

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

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

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

APPROXIMATE CONVERSIONS FROM SI UNITS

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

Integrating Non-Motorist Facility Data into Comprehensive Road Safety Assessment

A Center for Pedestrian and Bicyclist Safety Research Report

July 2024

Xiao Qin
Lawrence E. Sivak 71 Professor
Department of Civil and Environmental Engineering
Director, Institute for Physical Infrastructure and Transportation (IPIT)
University of Wisconsin-Milwaukee (UWM)

Robert Schneider
Professor and Chair
Department of Urban Planning
University of Wisconsin-Milwaukee (UWM)

Md Abu Sayed
Research Associate, IPIT
University of Wisconsin-Milwaukee (UWM)

Mohammad Abrari Vajari
Graduate Research Assistant
Department of Civil and Environmental Engineering
University of Wisconsin-Milwaukee (UWM)

TABLE OF CONTENTS

List of Figures

List of Tables

Acronyms, Abbreviations, and Symbols

VRUs	Vulnerable Road Users
FHWA	Federal Highway Administration
MIRE	Model Inventory of Roadway Elements
LLM	Large Language Model
CNN	Convolutional Neural Network
YOLO	You Only Look Once
CV	Computer Vision
GPT	Generative Pre-trained Transformer
GIS	Geographic Information System
API	Application Programming Interface
NLP	Natural Language Processing
ADA	Americans with Disabilities Act

Abstract

As active transportation gains popularity for both work and leisure trips, the need for reliable and accurate data on non-motorist activities and infrastructure becomes increasingly critical. Traditionally, there has been a lack of detailed data on non-motorist activities and facilities. Traditionally, detailed data on non-motorist activities and facilities has been scarce and challenging to obtain. However, advancements in technology now provide an opportunity to collect these data. This study aims to investigate recent technological advancements and methodologies for accessing non-motorist data sources and evaluate their ability to detect, classify, monitor changes in, and assess the condition of these facilities. A key focus is the use of vision-based large language models, specifically ChatGPT, for retrieving and evaluating non-motorist facilities from aerial images. This study tested ChatGPT's performance on a dataset containing satellite images to evaluate its ability to accurately detect and identify features such as pedestrian crosswalks. The results showed that ChatGPT can reliably assess crosswalk and sidewalk conditions, informing maintenance and improvement strategies. However, limitations were observed, including inconsistent detection of bicycle lanes without explicit visual cues and challenges in classifying crosswalks into specific types. Additionally, it struggles with reliably identifying street lighting and pedestrian signals, which are critical for comprehensive safety assessments.

Executive Summary

Addressing the surge in pedestrian and bicyclist fatalities requires integrating non-motorist facility data into comprehensive safety assessments. By employing new data collection methods, advanced technologies, and thorough safety analysis, researchers and practitioners can gain crucial insights into the roadway features that affect non-motorist safety. This understanding will inform strategies to improve traffic safety for pedestrians and bicyclists, ultimately aiming to reduce non-motorist fatalities.

To achieve this goal, this project includes the following tasks:

1. **Literature Review:** A comprehensive review was conducted, initially focusing on studies that discuss disparities in the accessibility of non-motorist safety facilities, considering geographic, demographic, and socioeconomic factors. This was followed by a review of the application of deep learning models and large language models (LLMs) such as ChatGPT in non-motorized transportation studies. The review identified significant gaps in maintaining an up-to-date inventory of non-motorist facilities and highlighted the potential of aerial images, computer vision, and ChatGPT to address these deficiencies.

2. **Data Source Exploration:** This study explored potential image data sources that could include non-motorist facility data and Google Earth Pro was identified as a primary resource for obtaining suitable images.

3. **ChatGPT Performance Assessment:** This study assessed the effectiveness of vision-based LLMs, specifically ChatGPT, in extracting information about non-motorist facilities. The evaluation involved comprehensive testing and analysis to determine the accuracy, reliability, and practicality of ChatGPT in identifying and evaluating features such as crosswalks, bicycle lanes, and other pedestrian infrastructures.

4. **Recommendations:** This study summarized the strengths and limitations of vision-based ChatGPT applications. It highlighted the challenges associated with using ChatGPT and aerial images for evaluating non-motorist safety facilities on roads.

This research is one of the first attempts to use vision-based LLMs, specifically ChatGPT, to extract non-motorist facility data from aerial images for assessments. The result underscores the importance of using vision-based LLMs like ChatGPT, customized for non-motorist data. By scaling this methodology, the accuracy and utility of these models in addressing issues related to infrastructure assessment and transportation safety for pedestrians and bicyclists can be significantly enhanced.

Introduction

In recent years, active transportation has gained popularity for both work and leisure trips. This trend surged during the COVID-19 pandemic, marked by a significant 57% increase in U.S. bicycle sales from April 2020 to April 2021(Francke, 2022). Despite this increasing growing interest, there has been a concerning rise in pedestrian and bicyclist fatalities since 2010, underscoring the need for effective policies and strategies. Previous research indicates that building and maintaining non-motorist facilities, along with ensuring their good condition, can increase both work and leisure-related non-motorized travel, as well as improve non-motorist safety. However, there is currently a notable disparity in the availability and quality of these facilities between different communities, such as affluent and disadvantaged communities, with the gap widening over time. This disparity highlights the need for equity in the distribution of non-motorist facilities for all communities, which can lead to improved protection for all members of society and promote inclusive, safe, and accessible transportation options.

Assessing and evaluating existing non-motorist facilities and infrastructure is essential to identify gaps in the network, prioritize non-motorist projects, maintain current facilities, and ensure they meet current demands and safety requirements, which requires updated and complete data on non-motorist activities and infrastructure inventories. However, despite guidelines for developing and maintaining these inventories, there is a noticeable lack of comprehensive data on non-motorized infrastructures, and the quality of existing inventories varies across different states.

As technology evolves, new and emerging techniques such as computer vision and large language models (LLMs) are revolutionizing large-scale data collection and analysis. For example, in active transportation, vision-based LLMs like ChatGPT can be used to effectively analyze and evaluate structured or unstructured data on non-motorist activities and facilities, providing additional insights and presenting results in a human-like manner, which was previously unattainable with traditional methods. Additionally, image processing technologies can be implemented to automatically analyze satellite images, identifying and assessing the condition of non-motorist facilities like crosswalks. Historical images can also be utilized to analyze past trends, identify effective interventions, and evaluate safety measures. The data derived from these technologies can pinpoint vulnerable communities, accident-prone areas, and risk factors, guiding targeted interventions, which is essential for developing predictive models, simulations, and optimizing resources. Moreover, these methods can aid in understanding the factors that drive the demand for pedestrian and bike facilities and in forecasting future needs.

By leveraging novel data collection methods, new data sources, advanced technologies, and innovative analytical approaches, including vision-based LLMs, this research project aims to:

- **Perform a Literature Review**: Investigate recent applications of artificial intelligence (AI) in non-motorist safety analysis.

- **Explore Image Data Sources**: Identify appropriate historical and current data sources at both intersection and area levels.

- **Assess ChatGPT Performance**: Evaluate ChatGPT's effectiveness in extracting information about non-motorist facilities. This task involves testing and analysis to determine the accuracy, reliability, and practicality of ChatGPT in identifying and evaluating features such as crosswalks, bicycle lanes, and other pedestrian infrastructures.

- **Summarize Findings**: Outline the strengths, limitations, and challenges associated with using image processing and LLMs to evaluate non-motorist facilities.

The outcomes of this research can serve as a foundation for evidence-based decision-making, policy development, and the design of targeted interventions to reduce pedestrian and bicyclist fatalities. By integrating the role of pedestrian and bicyclist facilities into the overall safety framework, transportation stakeholders can prioritize non-motorist safety in their planning, funding, and infrastructure improvement efforts. Ultimately, this research will contribute to creating more equitable and safe transportation systems that cater to the needs of all road users, ensuring that pedestrians and bicyclists can navigate roadways with reduced risk and increased confidence.

Literature Review

Every year, more than one million people lose their lives in traffic accidents globally, with nearly half of these fatalities involving non-motorists, also known as vulnerable road users (VRUs). (World Health Organization, 2015). With the growing popularity of walking and biking in recent years, an increase in fatalities among vulnerable road users is becoming a pressing concern (Das et al., 2021). The data from the National Highway Traffic Safety Administration reveals that in 2021, there were 7,388 pedestrian fatalities in road accidents in the United States, marking a 12.5% increase from the 6,565 pedestrian deaths in 2020, representing the highest number reported since 1981(National Center for Statistics and Analysis., 2023). Previous research shows that non-motorist crash patterns vary significantly depending on the location and type of road segment, with intersections being particularly hazardous for non-motorists (Hossain et al., 2023). Moreover, various studies have shown that underserved and low-income communities face an elevated risk of non-motorists injury and fatality due to the lack of adequate non-motorized infrastructure (Roll & McNeil, 2022). This emphasizes the urgent need to prioritize the safety of VRUs, particularly in high-risk areas like intersections and regions with low incomes and high poverty rates.

Previous research has explored the impact of various factors related to human, vehicle, environment, and roadway on non-motorists accidents. The findings from these studies emphasize that intersections present significant risks for VRUs (Haddad et al., 2023; Lee & Abdel-Aty, 2005). Intersections are specific road sections where the paths of vehicles and pedestrians traveling in different directions meet, leading to complex interactions and potential conflicts. This complexity is a critical factor contributing to many vehicle-pedestrian accidents. While various non-motorists safety facilities are developed and utilized at intersections to enhance pedestrian safety, one common approach involves implementing marked crosswalks, which serve as visible indicators of designated pedestrian crossing areas (Bian et al., 2020). As an example, a manual examination of satellite images revealed that approximately 60% of the nearly 6,400 intersections in San Francisco include crosswalks (Moran, 2022). The results of several studies provide evidence that crosswalks enhance pedestrian safety (Zegeer et al., 2005). Mitman et al. (2008) conducted a study to investigate pedestrian and driver behavior in locations with marked and unmarked crosswalks. The results revealed significant differences in how both drivers and pedestrians behave in these locations. Specifically, the findings demonstrated that drivers in areas with marked crosswalks are more inclined to yield to pedestrians compared to those in regions with unmarked crosswalks. Feldman et al. (2010) found a statistically significant reduction of approximately 40% in the number of accidents at intersections equipped with high-visibility crosswalks. It's important to note that the findings regarding the impact of crosswalks on pedestrian safety are not consistent across all studies. In contrast to the studies mentioned earlier, some research has indicated that unmarked crosswalks might actually have lower accident rates when compared to marked crosswalks (Bian et al., 2020). Anciaes and Jones (2018) and Ahmed et al. (2021) underscore the significant impact that the type, design, quality, and maintenance of crosswalks have on pedestrian preferences and experiences. Given the critical role of crosswalks in pedestrian safety, the FHWA recommends a comprehensive 6-step process for enhancing safety measures, which includes creating a crosswalk inventory map (Karaer et al., 2023).

Collecting crosswalk data and establishing a crosswalk inventory offer numerous benefits to transportation agencies, including gaining deeper insights into the causes of pedestrian-related accidents, identifying areas in need of additional crosswalks, exploring the potential for crosswalk removal or modification, and evaluating the overall pedestrian network connectivity (Twaddell et al., 2018). Traditional methods of crosswalk data collection such as field data collection and manual extraction from aerial images are prone to human errors, time-consuming, resource-intensive, and can disrupt traffic flow and raise safety concerns for data collectors (Y. Zhang et al., 2021). Moreover, the aging of materials in crosswalks underscores the inadequacy of a one-time data collection effort; thus, an ongoing crosswalk data collection procedure, at intervals of at least every several years, becomes crucial to consistently assess and address the evolving conditions of crosswalks (Chen et al., 2021). Currently, methods such as network buffering, collaborative mapping, and computer vision techniques have attracted attention in developing non-amortized inventories through the utilization of large-scale data (Karimi & Kasemsuppakorn, 2013).

Among these methods, computer vision techniques have emerged as the most promising, creating new opportunities for data collection, particularly in the context of crosswalk data (Huang et al., 2023). Computer vision (CV), an AI field, empowers computers and systems to extract valuable information from images and video data, enabling them to take actions or offer recommendations (Sharma et al., 2021). CV, which utilizes both machine learning (ML) and deep learning (DL) techniques, can be employed to recognize the presence, type, and location of objects in visual input. This technology finds a wide range of applications in fields such as healthcare, safety, retail, manufacturing, agriculture, and transportation (Kaya et al., 2023). Several studies have employed CV techniques for crosswalk detection using images. While some of these studies have shown good performance metrics, they do come with certain considerations. The majority of these investigations focused on Street View images (SVI) or images captured by cameras installed on or inside vehicles (e.g. dashcam) (Fan et al., 2020). Only a few studies have explored the utilization of remote sensing images (i.e., images taken from airplanes, drones, or satellites) (Chen et al., 2021). Although SVIs have recently become a crucial data source for geospatial data collection and urban analytics, they do possess inherent limitations that may impact their utility, especially for large-scale data collection efforts. These include limited geographic coverage, restricted temporal coverage with infrequent revisit periods (sometimes only once a decade), potential bias towards specific city areas, and privacy concerns (Gaw et al., 2022).

On the other hand, imagery obtained from satellite platforms, a vital data source for research in the built environment, provides broader spatial coverage, finer time granularity, and a relatively smaller data volume compared to alternative sources (Baučić et al., 2020). These characteristics make it a desirable choice for large-scale data collection and monitoring changes over time. However, detecting crosswalks from images, particularly satellite images, presents significant challenges, to name a few: complex backgrounds, including elements like rooftops or other road markings that may resemble crosswalks, the diverse environments and locations of crosswalks (e.g., urban, suburban, or varying weather conditions), distinct crosswalk designs, resolution constraints specific to satellite imagery, and potential interference from pedestrians and vehicles (Tümen & Ergen, 2020).

Given the complexities of the object detection task, several algorithms with diverse structures have been developed and employed for this purpose. In general, the evolution of object detection tasks can be divided into two primary stages: the first stage involves traditional object detection algorithms, while the second stage is marked by the adoption of detection models based on deep learning (Tian et al., 2021). In traditional approaches, researchers have employed simple image processing techniques such as edge detection, corner detection, or image segmentation to identify the presence of specific objects, like crosswalks, in an image. Traditional machine learning methods, including multilayer perceptron machines, support vector machines, and Random Forest classifiers, have also been applied for image classification tasks. These methods often use shallow structures to handle a limited number of samples and may face challenges in terms of performance and generalization abilities for complex classification and detection tasks (Xin & Wang, 2019).

Following the introduction of the Convolutional Neural Network (CNN) (Krizhevsky et al., 2012) in 2012 and Region Convolutional Neural Network (R-CNN) (Girshick et al., 2014) in 2014 DL-based algorithms have transformed image recognition and object detection tasks, outperforming traditional detection methods (Wu et al., 2020). These deep learning-driven object detection algorithms can be broadly categorized into two classes: one-stage detectors and two-stage detectors (D. Zhang et al., 2023). Two-stage detectors, characterized by their relatively more complex design, involve a two-step process of proposal generation and region classification. Common two-stage algorithms include Fast R-CNN, Faster R-CNN, Mask R-CNN, and Feature Pyramid Network (FPN). While two-stage detectors are more complex and typically deliver superior performance, their deployment on computationally constrained devices may be impractical (Carranza-García et al., 2021). In contrast, one-stage detection algorithms, such as You Only Look Once (YOLO) series, SqueezeDet, and Single Shot MultiBox Detector (SSD), streamline the process by eliminating a separate stage for proposal generation, providing efficient and precise detections in a single pass. For example, the latest version of YOLO series, namely YOLO v8, achieves high real-time performance and enhanced accuracy when executed on Graphics Processing Units (GPUs). As a result, the YOLO series stands out as one of the most common choices for detection tasks, owing to its simple structure, minimal complexity, and ease of implementation.

Several researchers have attempted to utilize and develop various YOLO models to detect various road regions and features, including crosswalks, from images. Kaya et al. employed both one-stage and two-stage algorithms, specifically the Faster R-CNN and YOLOv7 algorithms, for crosswalk detection (Kaya et al., 2023). Their dataset comprised 673 single-class images taken from the perspectives of both vehicles and pedestrians. The results highlighted the superior performance of the YOLOv7 model, achieving a mAP (mean average precision) of 98.6% at the GIoU threshold of 0.5, along with an F1-score value of 0.95. In another study, Karaer et al. developed a four-step procedure using aerial Images and the YOLOv2 model to create a crosswalk inventory map for the

State of Florida (Karaer et al., 2023). Initially, a GIS-based preprocessing approach was employed to select images intersecting with roadways. Subsequently, the YOLOv2 model was applied in the second step to detect crosswalks in the chosen aerial images. The third step involved an automatic mapping process to map all identified crosswalks using the YOLOv2 model. Finally, In the fourth step, performance evaluations were carried out by comparing the results with both a manually generated ground truth crosswalk dataset and an already existing crosswalk inventory. The results indicated that their framework achieved a recall of 85.9%, precision of 88.7%, and Intersection Over Union (IoU) of 76.9% on the ground truth dataset, demonstrating superior performance compared to the OpenStreetMap (OSM) dataset, which covered 77.8% of all crosswalks in the ground truth dataset. Verma et al. introduced a framework for automating the detection and inventory creation of pedestrian transportation infrastructure (Verma & Ukkusuri, 2023). Using the YOLOv5 large-size architecture, they identified crosswalks in publicly available satellite imagery and subsequently allocated the recognized crosswalks to intersection legs. The effectiveness of this approach was evaluated in Washington, D.C., and Los Angeles, CA, yielding classification accuracy rates of 71% and 89%, respectively.

As the field of computer vision progresses, the evolution of Large Language Models (LLMs) also improves our ability to analyze and understand extensive textual data. Recently, LLMs have attracted considerable interest due to their diverse applications across various sectors, including transportation (Zheng et al., 2023). A Large Language Model represents an advanced category of language models that utilizes deep learning techniques to process, generate, and understand human language on a vast scale. These models significantly outperform conventional statistical language models due to their complex architecture and extensive training on diverse datasets. Early language models were rule-based and limited in understanding context and the shift to statistical methods introduced probabilistic models like N-grams and Hidden Markov Models (Shoaib et al., 2023; Vaswani et al., 2017). Modern deep learning techniques, including RNNs, LSTMs, and Transformer-based models like BERT (Devlin et al., 2019), GPT-2 (Radford et al., 2019), and T5 (Raffel et al., 2020), significantly advanced the field. Amon them, GPT (Generative Pre-trained Transformer) is probably the most well-known thanks to the astronomical popularity of ChatGPT developed by OpenAI. GPT models, such as GPT-3 and GPT-4, are designed to understand and generate human-like text based on the input they receive. They are built using a large number of parameters and trained on diverse datasets to perform various language tasks, including translation, summarization, question answering, and more. Scaling up training data and parameters led to GPT-3 (Brown et al., 2020) with 175 billion parameters, spurring the development of other models like Megatron-Turing NLG (Smith et al., 2022), Chinchilla (Hoffmann et al., 2022), PaLM (Chowdhery et al., 2023), OPT (S. Zhang et al., 2022), and LLaMA (Touvron et al., 2023). Recently, open-source models like Alpaca (Taori et al., 2023) and Vicuna (*Vicuna*, n.d.), developed from LLaMA, have shown comparable performance.

Computer vision and LLMs have been incredibly useful for tackling specific tasks predefined by humans. However, the ultimate goal has always been to develop systems capable of handling a wide range of problems with minimal adjustments from us. This vision extends to creating models that can describe visual inputs in human language, paving the way for visual LLMs or LLMs with vision capabilities, which naturally offer more versatility than pure computer vision systems (Ju et al., 2022).

In conclusion, previous research has shown that there is currently a lack of data related to non-motorist facilities and infrastructure, which can hinder the development of safety measures for non-motorists. However, the advancement of technologies such as large language models (LLMs) and computer vision (CV) provides new methods for collecting, analyzing, and evaluating data on non-motorist activities and facilities. These innovative approaches offer the potential to fill existing data gaps, enhance the understanding of non-motorist safety issues, and guide targeted interventions. By leveraging these advanced technologies, it is possible to improve the overall safety and accessibility of non-motorist infrastructure, ultimately benefiting all members of society.

Methodology

In the field of vision-based large language models (LLMs), several advanced deep learning models have been successfully tested in various domains. This project explored these models with a focus on ChatGPT and leveraged their capabilities to develop pedestrian and bicyclist facility detection and safety assessment models.

ChatGPT-4o

In this research, the GPT-4o through Application Programming Interface (API) was chosen over its online version due to the greater control on the dynamic nature of ChatGPT's responses. Through the API, parameters can be fine-tuned, and specific constraints set to ensure that the output aligns precisely with defined requirements. This level of customization is less attainable with the standard online interface. Additionally, leveraging the API allows for seamless integration of GPT-4o into various other applications and workflows that enhance functionality and efficiency. The API also facilitates the automation of tasks such as data processing, content generation, and customer support, significantly reducing manual efforts and operational costs. Moreover, the API enables handling higher volumes of requests and data, ensuring that the system remains responsive and efficient even during peak usage times. This scalability is crucial for businesses and developers who rely on consistent performance. The API provides access to more granular analytics and usage metrics, enabling better monitoring and optimization of the system's performance. Overall, the GPT-4o API offers a robust, scalable, and customizable solution that enhances the versatility and effectiveness of ChatGPT that it a superior choice for advanced applications and integration.

In order to access the GPT through API, a tool called "AI image Analyzer" was developed that mirrors the capabilities of the online version but offers additional flexibility and adaptability tailored to specific needs. It can work with multiple images and receive user inputs similar to the GPT online version and generates responses as presented in Figure 1. In the developed tool, the dynamic nature of the response has been controlled by a parameter called temperature. Temperature[1] controls the "creativity" or randomness of the text generated by GPT. A higher temperature (e.g., 0.7) results in more diverse and creative output, allowing for a wider range of possible tokens at each step, which makes the text more varied. In contrast, a lower temperature (e.g., 0.2) makes the output more deterministic and focused, reducing variability and leading to more predictable and consistent text. A temperature of 0 makes the model completely deterministic, always choosing the most likely token based on the context. This tool can be embedded into websites, mobile applications, or other platforms, providing a seamless user experience.

[1] Cheat Sheet: Mastering Temperature and Top_p in ChatGPT API https://community.openai.com/t/cheat-sheet-mastering-temperature-and-top-p-in-chatgpt-api/172683

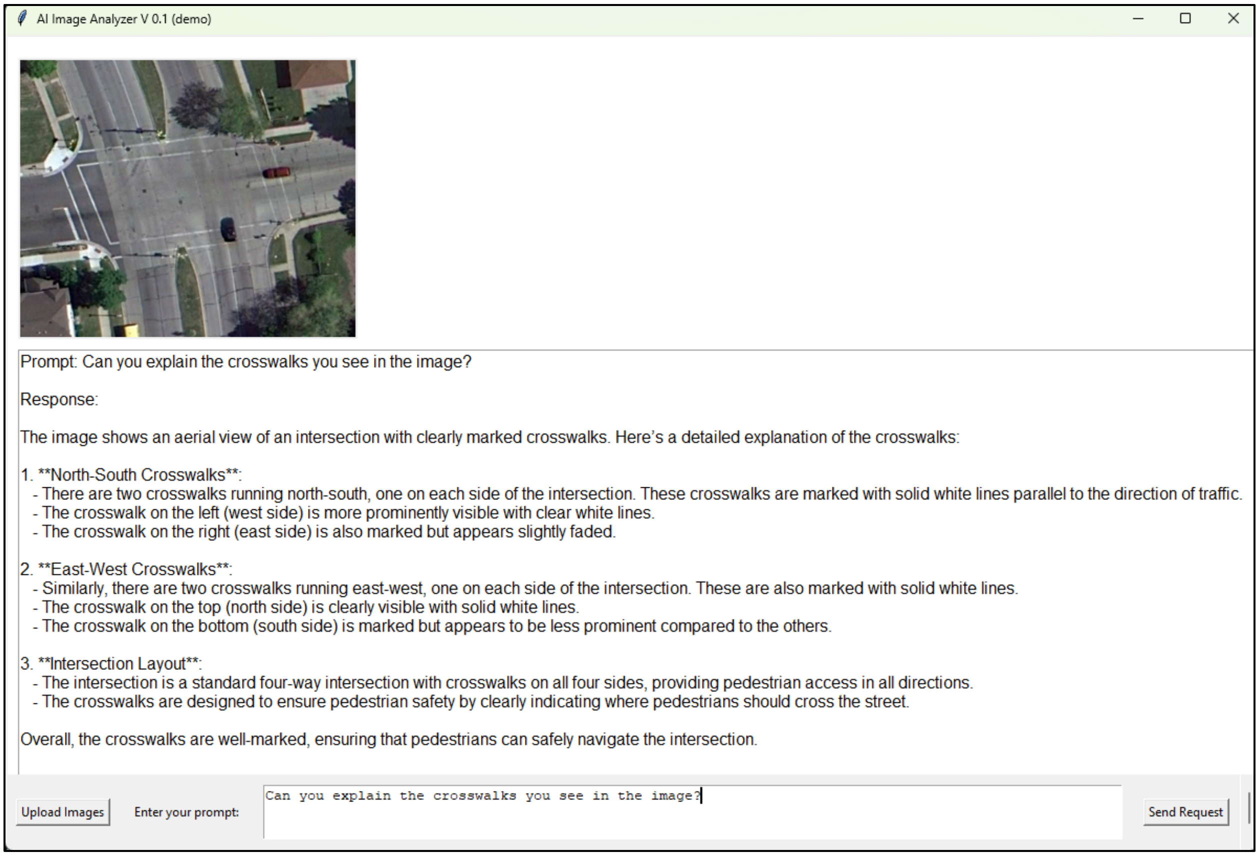


Figure 1. Interface Developed for ChatGPT

Prompt Design

Prompt writing, also termed prompt engineering, is pivotal in optimizing the performance of LLMs like GPT. It involves designing precise sets of instructions that tailor an LLM's behavior by customizing or refining its capabilities, serving as the primary interface for human interaction with these sophisticated natural language processing (NLP) systems (Arora et al., 2022; van Dis et al., 2023; White et al., 2023; Zhou et al., 2022). A prompt, in this context, refers to a predefined text or set of instructions provided to an LLM that guides its responses and behavior. This method has gained widespread acceptance due to its versatility and effectiveness across various NLP tasks. However, the efficacy of prompt engineering hinges significantly on the quality of the prompts themselves, which are meticulously designed to ensure they effectively steer the model's responses in desired directions. Unlike traditional methods for analyzing models, prompt-based approaches offer a non-invasive means of assessing an LLM's comprehension without requiring direct inspection of its internal mechanisms (Shin et al., 2020). This method provides a valuable baseline for evaluating the model's knowledge and performance, making it a valuable tool for researchers and developers alike. Nevertheless, manual prompt construction remains a labor-intensive process that demands meticulous attention to detail, particularly for tasks that require nuanced contextual understanding, such as textual entailment. Moreover, the sensitivity of LLMs to prompt construction underscores the challenge of achieving consistent high performance across diverse applications. Researchers are exploring automated techniques to streamline prompt generation, aiming to enhance efficiency and reliability while preserving the integrity of interactions between humans and AI systems (Arora et al., 2022; Jung et al., 2022; Liu et al., 2023; Reynolds & McDonell, 2021; Shin et al., 2020; Wei, Tay, et al., 2022; Wei, Wang, et al., 2022). This research intends to explore and evaluate the ability of a LLM with vision capabilities (ChatGPT-4o) to detect non-motorist facilities, identify issues, and generate insightful recommendations.

To achieve this objective, this research included various non-motorist facilities and then designed the prompt. Using the list, a series of objective questions has been developed for intersection level image analysis. Additionally, another set of subjective questions has been developed for both intersection and area level image analysis. During prompt design, multiple versions of prompts have been tested and the prompts that provided the most reasonable output have been selected as final prompts. Table 1 and Table 2 present summary of all the prompts used in this research for the evaluation of ChatGPT.

Table 1. Objective Questions to Detect and Assess Pedestrian and Bicycle Facilities

<table>
<tr><th colspan="2">Summary of Prompts for Objective Questions</th></tr>
<tr><td><u>Pedestrian crosswalk</u>
1. Can you see the pedestrian crosswalk from the picture?
2. Where is the pedestrian crosswalk located in the picture?
3. Is the crosswalk clearly marked?
4. Can you describe the condition of the crosswalk (e.g. good, fair, poor)?
5. Are the sidewalks accessible?
6. Is the crosswalk accompanied by any traffic signals or signage?
7. Are there any safety features associated with the crosswalk?
8. Can you see the pedestrian sidewalks from the picture?</td><td><u>Bicycle lane</u>
1. Where is the bicycle lane located in the picture?
2. Is the bicycle lane clearly marked and unobstructed?
3. Can you describe the condition of the bicycle lane (e.g. good, fair, poor)?
4. Is the bicycle lane accessible?
5. Does the lane have any safety features like buffers from traffic or signage?
6. Can you see the ADA curb ramps from the picture?</td></tr>
<tr><td><u>Pedestrian sidewalks</u>
1. Where are the pedestrian sidewalks located in the picture?
2. Are the sidewalks clearly marked?
3. Can you describe the condition of the sidewalks (e.g. good, fair, poor)?
4. Are the sidewalks accessible?
5. Do the sidewalks have any features like benches, trees, or lampposts?
6. Can you see the bicycle lane from the picture?</td><td><u>Pedestrian refuge areas</u>
1. Where are the pedestrian refuge areas located in the picture?
2. Are they adequately sized and positioned for pedestrian safety?
3. Can you describe the condition of the pedestrian refuge areas (e.g. good, fair, poor)?
4. Do they include any features like signage or lighting?</td></tr>
<tr><td><u>On-street parking</u>
1. Can you see the on-street parking from the picture?
2. Where is the on-street parking located in the picture?
3. Is the parking clearly marked and organized?
4. Can you describe the condition of the parking (e.g., good, fair, poor)?
5. Are there any safety features like designated parking spots for disabled individuals?</td><td><u>ADA curb ramps</u>
1. Where are the ADA curb ramps located in the picture?
2. Are the ramps compliant with ADA standards?
3. Can you describe the condition of the ADA curb ramps (e.g. good, fair, poor)?
4. Do the ramps have detectable warning surfaces for visually impaired individuals?
5. Can you see the curb extensions or ‘bulb-outs’ from the picture?</td></tr>
<tr><td colspan="2"><u>Curb extensions</u>
1. Where are the curb extensions located in the picture?
2. What impact do the curb extensions appear to have on traffic and pedestrian safety?
3. Can you describe the condition of the curb extensions or ‘bulb-outs’ (e.g., good, fair, poor)?
4. Can you see the pedestrian refuge areas from the picture?</td></tr>
</table>

Table 2. Subjective Questions for the Evaluation of ped/bike safety and walkability

Summary of Prompts for Intersection level analysis
1. How safe is it for a pedestrian to cross any leg of the intersection? 2. Describe the safety conditions of a pedestrian crossing any leg of the intersection. 3. Are there any safety concerns for pedestrians crossing any leg of the intersection? If yes, what are they? 4. Can you provide any recommendations to improve safe crossing for pedestrians? 5. These two images are of the same intersection. The first one is from 2010 and the second one is from 2021. 6. What could be potential reasons for the crosswalk not being present in the picture? 7. Are there any other pedestrian facilities visible in the picture? 8. What might be the reasons for the absence of sidewalks in this intersection? 9. Are there any other pedestrian facilities visible in the picture? 10. Could there be another facility for cyclists in the area? 11. Are there any indicators that a bicycle lane might be needed or planned? 12. What could be potential reasons for the absence of ADA curb ramps? 13. Are there other accessibility features visible in the picture? 14. What might be the benefits of adding curb extensions in this area? 15. Are there any other traffic calming measures visible? 16. What might be the reasons for the absence of pedestrian refuge areas? 17. Are there any other safety measures for pedestrians visible in the picture? 18. These two images are of the same intersection. The first one is from 2010 and the second one is from 2021. - Were there any improvements to pedestrian safety while crossing any leg of the intersection? If so, what are they? - Were there any aspects of pedestrian safety that deteriorated or got worse from 2010 to 2021?
Summary of Prompts for Area level Analysis
<u>Walkability</u> 1. Please rate the picture for walkability in this neighborhood. Describe the walking conditions you observe, such as the presence and quality of sidewalks, crosswalks, pedestrian signals, street lighting, and overall pedestrian safety. 2. Please provide a walk score for the neighborhood shown in the picture. Describe the features that contribute to your score, such as sidewalks, crosswalks, pedestrian signals, street lighting, and overall pedestrian safety. **<u>Bikeability:</u>** 1. Please rate the picture for bikeability in this neighborhood. Describe the biking conditions you observe, including the presence and quality of bicycle lanes, bike racks, traffic conditions, street lighting, and overall safety for cyclists. 2. Please provide a bike score for the neighborhood shown in the picture. Describe the features that contribute to your score, such as bicycle lanes, bike racks, traffic conditions, street lighting, and overall safety for cyclists.

Data Collection

In order to evaluate the performance of the ChatGPT model in detecting non-motorists facilities in the images, a dataset containing images with appropriate dimension and resolution was required. Initially, various data sources including the Google Maps platform, Google Earth Pro, and Bing Maps were explored to select the best source for collecting images. Factors considered included the quality of images, availability of historical data, the availability of different zoom levels and views, frequency of image updates, geographic coverage, and the ease of integration with our API. After evaluating these comprehensive factors, Google Earth Pro was selected as the optimal source. The satellite images showcasing various non-motorist facilities for selected locations were then retrieved from it to test the GPT model. Below is a description of the different facilities considered in the analysis:

Non-motorist Facilities

Pedestrian crosswalks are essential for providing safe crossing points across streets, and evaluating their presence, condition, and visibility is necessary to ensure they are clearly marked, well-maintained, and supported by appropriate signage and traffic signals. Properly designed crosswalks not only help prevent accidents and enhance pedestrian safety but also should be clearly marked with visible paint, complemented by traffic signals and pedestrian signs, and regularly maintained to prevent wear. These features, identifiable in aerial images through their distinct markings and strategic positioning at intersections, play a critical role in ensuring pedestrian safety and effective traffic management.

Sidewalks offer designated pathways for pedestrians, separating them from vehicular traffic. Evaluating sidewalks for clear markings, good condition, and accessibility is crucial for the safety and mobility of all pedestrians, including those with disabilities. Sidewalks must be free of cracks and obstructions, sufficiently wide to accommodate wheelchairs and strollers, and include amenities like benches, trees, and lampposts for enhanced pedestrian comfort. These elements can be observed in aerial images by examining the pathways running alongside roads and their overall continuity and condition.

Bicycle lanes provide cyclists with dedicated space, reducing the risk of collisions with motor vehicles. Evaluating the condition and accessibility of these lanes is essential for cyclist safety and promoting bicycle use as a mode of transportation. It should be clearly marked with paint and signage, kept unobstructed and in good condition, and include safety features such as buffers from traffic and reflective markings. Aerial images can reveal these lanes through their distinct markings parallel to roadways, highlighting their connectivity and condition.

ADA curb ramps are vital for ensuring accessibility for individuals with disabilities. Assessing these ramps for compliance with ADA standards ensures that all pedestrians can navigate the urban environment safely and independently. These ramps should be installed at all crosswalks, maintained free of damage and obstructions, and feature detectable warning surfaces for the visually impaired. While detailed compliance with ADA standards may not be visible, the presence and location of curb ramps can be confirmed through aerial images at intersections.

Curb extensions, or bulb-outs, reduce the crossing distance for pedestrians, enhance visibility, and act as traffic calming measures. Evaluating their impact on traffic and pedestrian safety helps design safer intersections. They should be installed at high-traffic intersections, maintained in good condition, and regularly evaluated for their impact on traffic flow and pedestrian safety. Curb extensions can be identified in aerial images by their protruding structures at intersections.

Pedestrian refuge areas provide safe waiting spaces for pedestrians crossing wide or busy streets. Assessing these areas ensures they are adequately sized and strategically positioned for optimal safety. Pedestrian refuge areas should be constructed at busy intersections, equipped with signage and lighting, and maintained in good condition. These areas can be spotted in aerial images as median islands or other protected spaces within crosswalks.

On-street parking can impact pedestrian and cyclist safety by affecting visibility and traffic flow. Evaluating its organization and condition is essential to ensure it does not pose safety hazards. It should be clearly marked, include designated spots for disabled individuals, and be regularly maintained to prevent damage and obstructions. The layout and organization of on-street parking can be effectively assessed through aerial views, providing insights into its impact on urban mobility and safety.

Facility Locations

This research initially used 10 intersection level and 2 area level images (Figure 2) for the test.

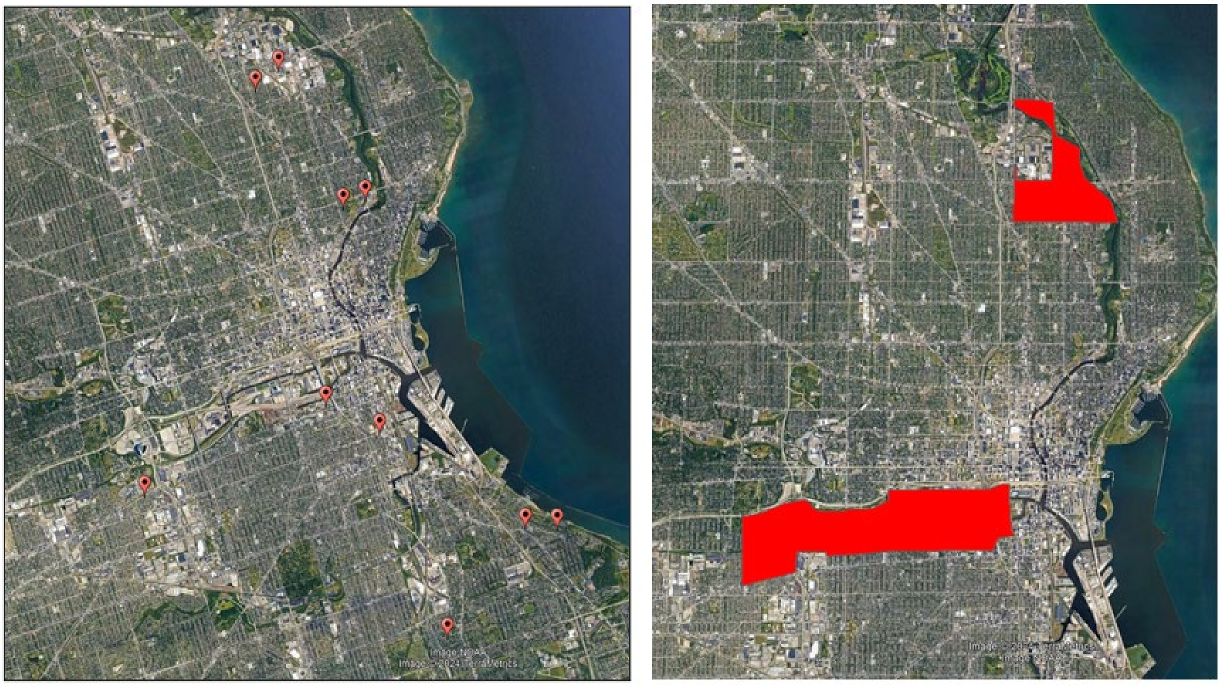


Figure 2. Location of Selected Intersections and Zones for ChatGPT Analysis

The evaluation is divided into two parts: intersection-level analysis and area-level analysis. For the intersection-level analysis, the focus was on the safety and accessibility of pedestrian crossings, bicycle lanes, and related infrastructure. Prompts were designed to extract detailed observations about the condition, visibility, and safety features of these facilities. For the area-level analysis, the overall walkability and bikeability of neighborhoods were assessed, taking into account factors such as the presence and quality of sidewalks, crosswalks, bicycle lanes, pedestrian signals, and street lighting.

Results

This section presents the findings from the evaluation of the effectiveness and condition of non-motorist facilities using image analysis prompts developed for both intersection and area levels. The non-motorist infrastructures include pedestrian crosswalks, bicycle lanes, sidewalks, pedestrian refuge areas, on-street parking, ADA curb ramps, and curb extensions. By employing a series of carefully designed prompts, this research utilized ChatGPT to analyze images and provide objective and subjective insights into the current state and safety of these facilities.

The evaluation is performed at intersection-level and area-level analysis, respectively. For the intersection-level analysis, the focus was on the safety and accessibility of pedestrian crossings, bicycle lanes, and related infrastructure. Prompts were designed to extract observations about the condition, visibility, and safety features of these facilities. For the area-level analysis, the overall walkability and bikeability of neighborhoods were assessed, considering factors such as the presence and quality of sidewalks, crosswalks, bicycle lanes, pedestrian signals, and street lighting.

Non-motorist Feature Detection

Pedestrian Crosswalks

1. ChatGPT was able to detect crosswalks, but consistently classify whether they are traverse, ladder or zebra crossing.
2. It was able to explain the location of the crosswalks, but used various terminology such as "north, south, east and west", "all four corners of the intersection", and "each corner of the intersection".
3. The response was consistent for evaluating crosswalk conditions.
4. The result was not consistent in detecting other features such as signage and lighting.

Bicycle Lanes

The results were inconsistent for detecting bicycle lanes. ChatGPT was looking for clear signage or bike-image in the images.

Pedestrian Sidewalks and Refuge Areas

1. ChatGPT was able to detect this feature.
2. It was able to explain the location of the features, but used various terminology such as "north, south, east and west", "all four corners of the intersection" and "each corner of the intersection".
3. The response was consistent for evaluating its conditions.

On-Street Parking and ADA Curb Ramps

ChatGPT was able to detect this feature but unable to evaluate the condition of this feature. It properly explained why using ariel view image this cannot be done.

Curb Extensions

1. ChatGPT was able to detect this feature, but did not consistently classify whether they are traverse, ladder or zebra crossing.
2. It was able to explain the location of the crosswalks, but used various terminology such as "north, south, east and west", "all four corners of the intersection", and "each corner of the intersection".
3. The response was consistent for evaluating the conditions.

Walkability and Bikeability

To assess the walkability and Bikability of a neighborhood based on aerial images, the following features and evaluation strategy are considered by ChatGPT:

1. Sidewalks: Presence and condition of sidewalks.
2. Crosswalks: Marking condition.
3. Street Lighting: ChatGPT looked for this image but could not detect it.
4. Pedestrian Signals: it could not detect but provided good reasoning suggesting that this feature was likely to be present. It counted that the pedestrian signal was present.
5. Overall Pedestrian Safety: It used neighborhood design pattern, presence of green spaces and trees to explain overall pedestrian safety.

Based on the evaluation of the above features, its rating on walkability of the neighborhood was reasonable.

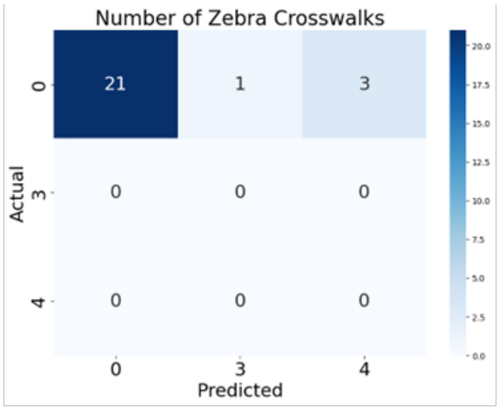
Performance of ChatGPT

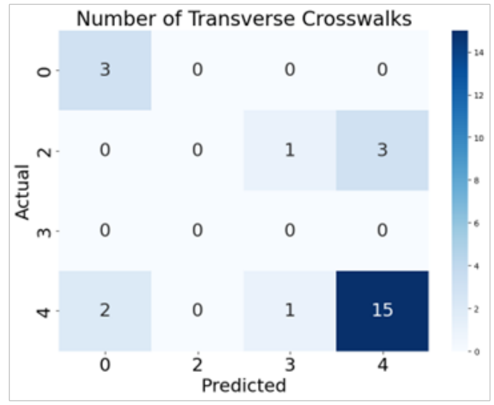
This study used confusion matrix and accuracy to evaluate the performance of ChatGPT in different detection tasks. Utilizing both overall and per-class accuracy metrics derived from the confusion matrix provides a holistic view of a classification model's performance. This analysis is crucial for optimizing models to achieve balanced performance across all classes, ensuring that the model is both accurate and reliable in real-world applications. A confusion matrix is an essential tool for evaluating the performance of a classification model on a dataset where the true values are known. It is particularly valuable in multiclass classification settings, as it provides a comprehensive overview of how well the model predicts across multiple classes.

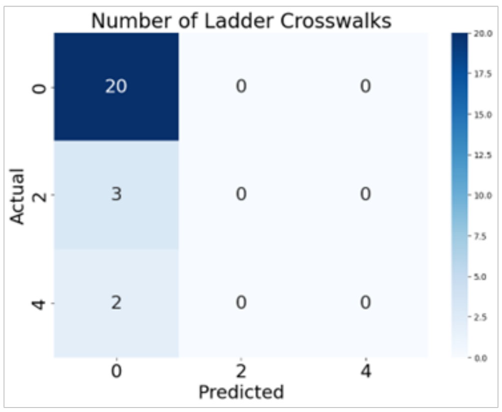
The overall accuracy of a model derived from the confusion matrix is calculated by the ratio of the sum of the diagonal elements (correct predictions) to the total number of predictions made.

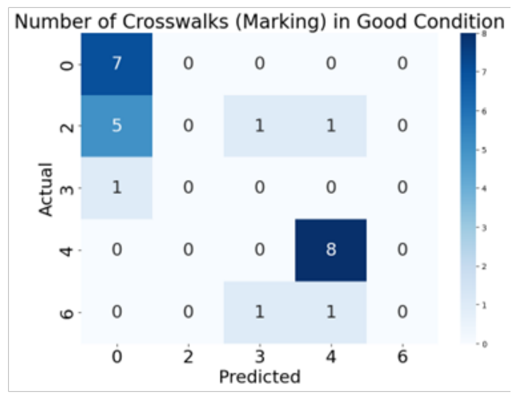
The formula is expressed as: Overall Accuracy = (Sum of correct predictions (diagonal sum)) / (Total number of predictions (all matrix elements sum))

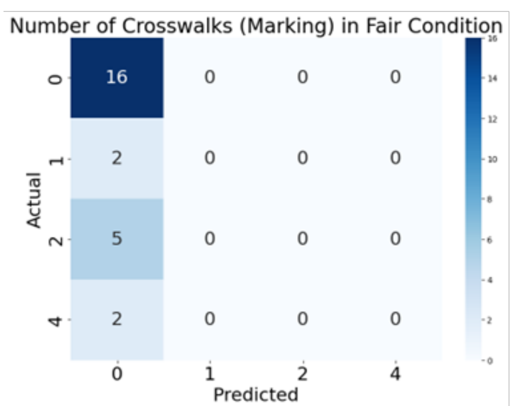
This metric provides a quick snapshot of the model's effectiveness across all classes. However, it might not fully represent the model's performance in cases of class imbalance, highlighting the need for accuracy metrics specific to each class.

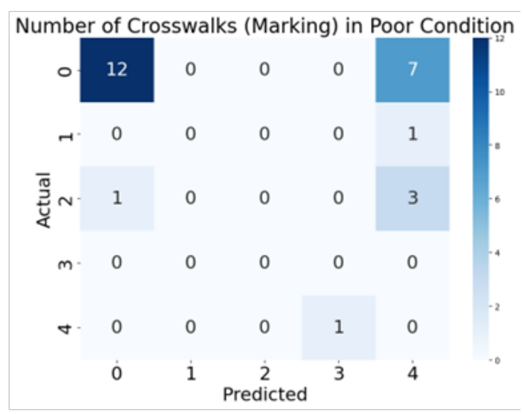

Number of Zebra Crosswalks
Actual
Predicted
0
3
4
21
1
3
0
0
0
0
0
0

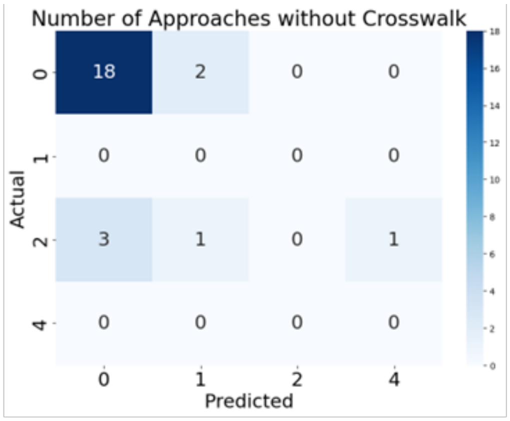

Number of Transverse Crosswalks
Actual
Predicted
0
2
3
4
3
0
0
0
0
0
1
3
0
0
0
0
2
0
1
15

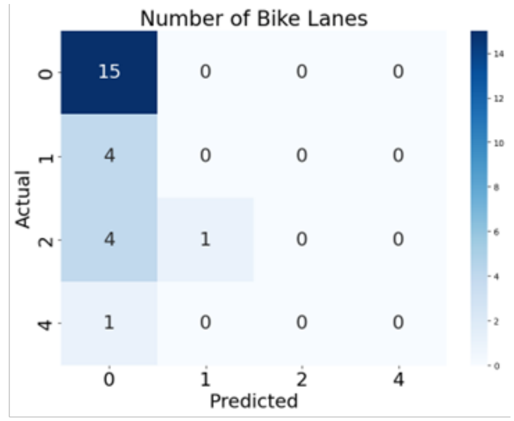

Number of Ladder Crosswalks
Actual
Predicted
0
2
4
20
0
0
3
0
0
2
0
0

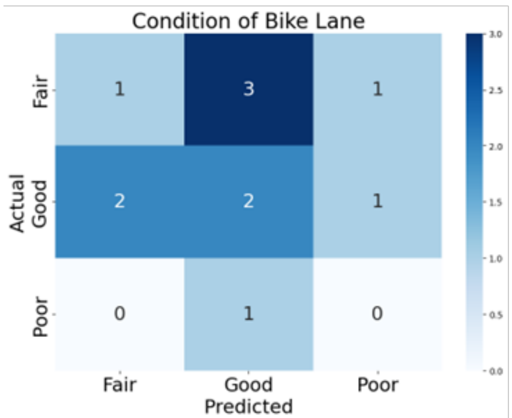

Number of Crosswalks (Marking) in Good Condition
Actual
Predicted
0
2
3
4
6
7
0
0
0
0
5
0
1
1
0
1
0
0
0
0
0
0
0
8
0
0
0
1
1
0

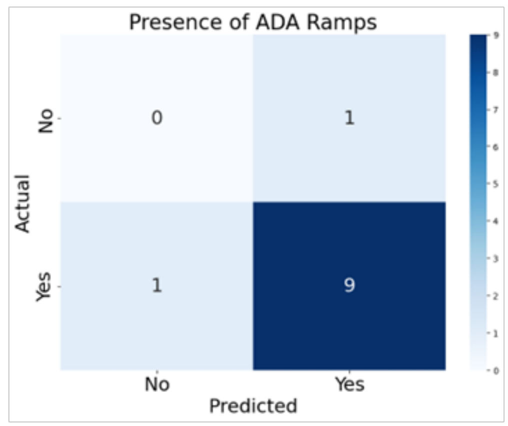

Number of Crosswalks (Marking) in Fair Condition
Actual
Predicted
0
1
2
4
16
0
0
0
2
0
0
0
5
0
0
0
2
0
0
0

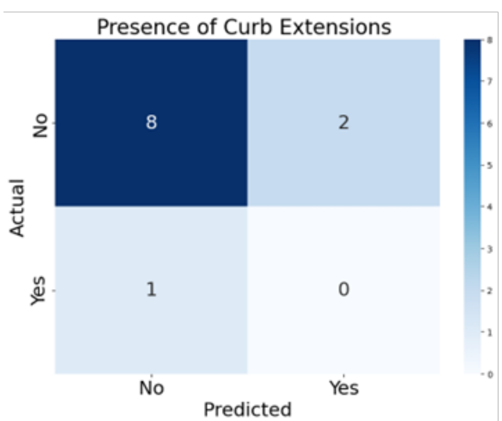

Number of Crosswalks (Marking) in Poor Condition
Actual
Predicted
0
1
2
3
4
12
0
0
0
7
0
0
0
0
1
1
0
0
0
3
0
0
0
0
0
0
0
0
1
0

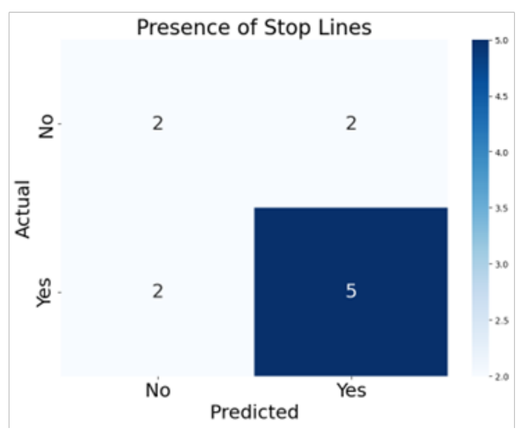

Number of Approaches without Crosswalk
Actual
Predicted
0
1
2
4
18
2
0
0
0
0
0
0
3
1
0
1
0
0
0
0


Number of Bike Lanes
Actual
Predicted
0
1
2
4
15
0
0
0
4
0
0
0
4
1
0
0
1
0
0
0


Condition of Bike Lane
Actual
Predicted
Fair
Good
Poor
1
3
1
2
2
1
0
1
0


Presence of ADA Ramps
Actual
Predicted
No
Yes
0
1
1
9


Presence of Curb Extensions
Actual
Predicted
No
Yes
8
2
1
0


Presence of Stop Lines
Actual
Predicted
No
Yes
2
2
2
5

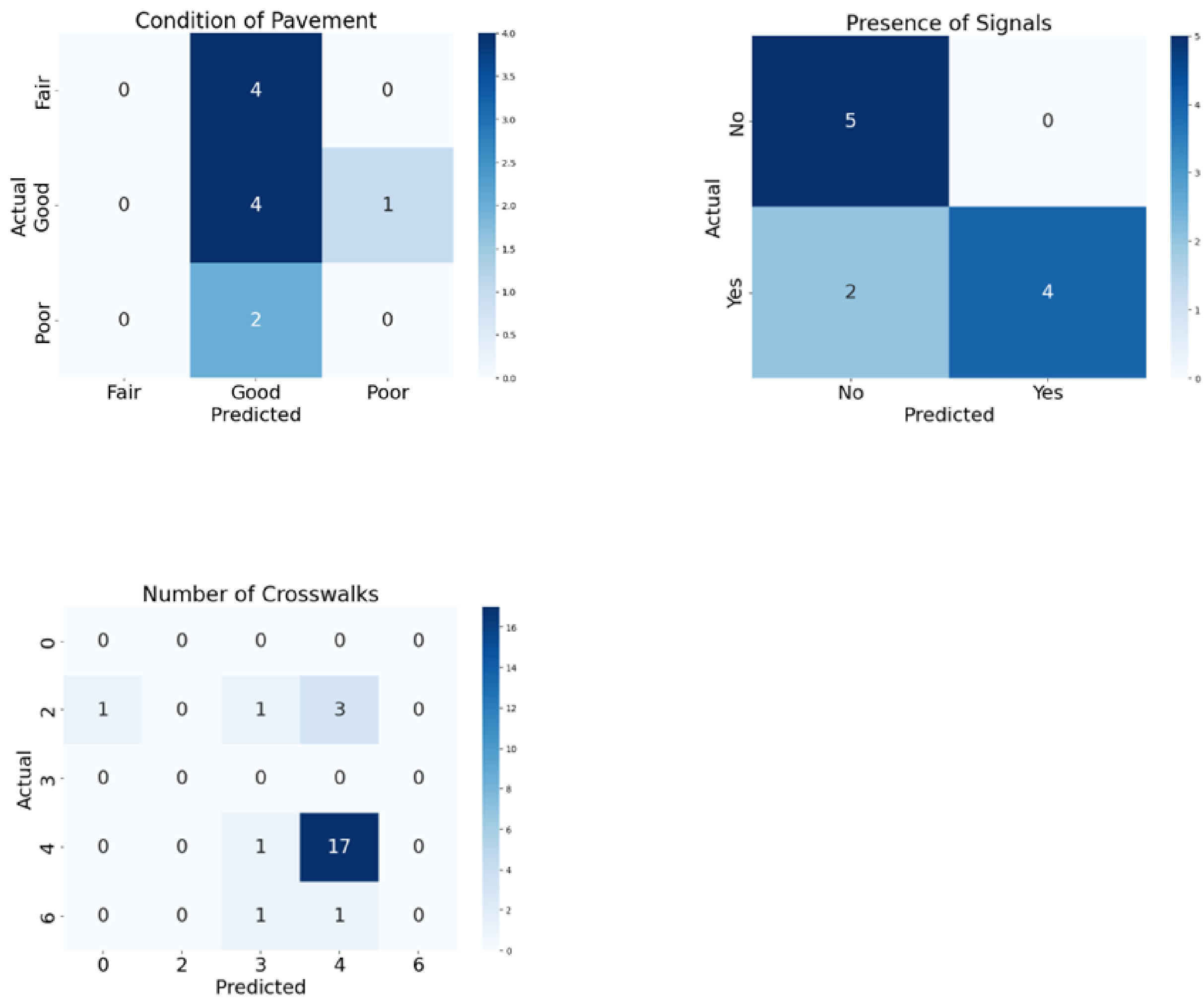


Figure 3. Confusion Matrices for Non-motorist Facilities Detection by ChatGPT

Per-class accuracy measures the effectiveness of the model in predicting each class individually. It is calculated by dividing the number of correct predictions for a specific class (a diagonal element) by the total number of instances that actually belong to that class (the sum of elements in the corresponding row). This measure is vital for pinpointing which classes are predicted with high precision and which are not, guiding targeted improvements in model training. The formula for this measure is as follows:

Per-Class Accuracy for Class i = (Correct predictions for Class i) / (Total actual occurrences of Class i).

This measure is vital for pinpointing which classes are predicted with high precision and which are not, guiding targeted improvements in model training.

The following Figure 3 presents confusion matrices for the detection and assessment of non-motorist facilities by ChatGPT. For each prompt, multiple runs at different times were conducted to ensure the consistency of the predictions.

Table 3 also shows the accuracy of ChatGPT in detecting various non-motorist facilities in images, such as crosswalks, their conditions, and bicycle lanes. This table presents the overall accuracy of ChatGPT in detecting each non-motorist facility, along with the per-class accuracy for each facility. The evaluation of various pedestrian and bicycle facilities indicates differing levels of detection accuracy. The ability to count crosswalks achieved moderate accuracy at 0.68, suggesting a reasonable capability. For crosswalk markings, the detection accuracy varied with condition: markings in good condition had a moderate accuracy of 0.60, fair condition markings achieved 0.64, while poor condition markings had lower accuracy at 0.48, indicating difficulty in detection. Approaches without crosswalks showed good detection accuracy at 0.72. Zebra crosswalks and ladder crosswalks both demonstrated high detection accuracy at 0.84 and 0.80 respectively, indicating strong detection capabilities. Transverse crosswalks also had good accuracy at 0.72, suggesting reliable detection. Lastly, the detection of bicycle lanes was perfect, with an accuracy of 1.00, indicating excellent detection capability.

Table 3. Performance of ChatGPT Detecting and Counting Non-motorist Facilities

	Accuracy								
Classes	Number of Crosswalks	Number of Crosswalks (Marking) in Good Condition	Number of Crosswalks (Marking) in Fair Condition	Number of Crosswalks (Marking) in Poor Condition	Number of Approaches without Crosswalk	Number of Zebra Crosswalks	Number of Transverse Crosswalks	Number of Ladder Crosswalks	Number of Bicycle lanes
Overall	0.68	0.6	0.64	0.48	0.72	0.84	0.72	0.8	1
Class 0	-	1	1	0.63	0.9	0.84	1	1	1
Class 1	-	-	0	0	-	-	0	-	0
Class 2	0	0	0	0	0	-	-	0	0
Class 3	-	0	-	-	-	-	-	-	-
Class 4	0.94	1	0	0	-	-	0.83	0	0
Class 5	-	0	-	-	-	-	-	-	-
Class 6	0	-	-	-	-	-	-	-	-

Class-specific detection accuracy varies significantly across different classes. When there are no crosswalks present, detection accuracy is high across most features, especially for crosswalks in good and fair condition, zebra crosswalks, transverse crosswalks, ladder crosswalks, and bicycle lanes. In class with 1 crosswalk, there is no detection for most features, indicating potential issues or lack of data for this case. When there are 2 crosswalks, there is no detection for all features, suggesting this scenario either lacks sufficient data or the detection algorithm is ineffective. For class with 3 crosswalks, there is limited data, and only poor accuracy is observed for crosswalks in good condition. When there are 4 crosswalks, detection accuracy is high for crosswalks in good condition and the overall number of crosswalks, though there are some issues with detecting crosswalks in fair and poor conditions. In scenarios with 5 and 6 crosswalks, there is minimal data or detection, indicating potential gaps in the dataset or challenges in detection for these scenarios.

Table 4 shows the detection accuracy for the conditions of bicycle lanes and pavements. The overall detection accuracy of the model indicates significant challenges in assessing the condition of bicycle lanes and pavements. Specifically, the system shows a low overall accuracy of 0.27 for bicycle lane conditions and a slightly better, yet still low, accuracy of 0.36 for pavement conditions, suggesting substantial struggles in evaluating these features accurately.

Table 4. Performance of ChatGPT in Assessing the Conditions of Non-motorist Facilities

		Accuracy	
	Classes	Condition of Bicycle Lane	Condition of Pavement
Condition Assessment	Overall	0.27	0.36
	Fair	0.2	0
	Good	0.4	0.8
	Poor	0	0

For bicycle lanes in fair condition, the accuracy is very low at 0.2, indicating difficulties in reliable detection, while there is no detection capability (0.0) for pavements in fair condition, suggesting a complete failure in recognition. When it comes to detecting bicycle lanes in good condition, the system exhibits moderate accuracy (0.4), showing some capability. The model performs well in identifying pavements in good condition, with a high accuracy of 0.8. However, the detection of bicycle lanes and pavements in poor condition is non-existent (0.0), highlighting a significant gap in the model's abilities.

Table 5 presents findings for assessing how well ChatGPT performs in identifying different non-motorist facilities from aerial photographs. The table summarizes the system's accuracy in detecting the presence or absence of curb extensions, stop lines, ADA ramps, and signals.

Table 5. Performance of ChatGPT in Detecting Other Non-Motorist Facilities

Detection	Classes	Accuracy			
		Presence of Curb Extensions	Presence of Stop Lines	Presence of ADA Ramps	Presence of Signals
	Overall	0.72	0.64	0.81	0.81
	Yes	0	0.71	0.9	0.68
	No	0.8	0.5	0	1

The GPT model demonstrates strong accuracy in detecting ADA ramps and traffic signals, both achieving an overall accuracy of 0.81. This suggests robust capabilities in recognizing these crucial non-motorist facilities. In contrast, the detection of curb extensions shows moderately good performance with an accuracy of 0.72, indicating effective identification in most cases but with room for improvement. The lowest accuracy is observed in the detection of stop lines, with an accuracy of only 0.64, highlighting challenges in reliably identifying these important road markings.

Overall, while ChatGPT shows potential in certain areas, especially in identifying well-maintained pavements, it struggles significantly in other areas, particularly in recognizing fair and poorly maintained conditions. While the GPT excels in detecting certain non-motorist features like signals and ADA ramps, it faces challenges in consistently identifying stop lines and effectively detecting ADA ramps when absent.

Preliminary Application: Changes in Non-motorist Facilities at a Sample of Milwaukee Intersections: 2011 vs. 2022

While the detection capabilities of ChatGPT are still limited and require further development, this section describes the application of this tool in 30 randomly selected census tracts in the City of Milwaukee, WI (Figure 4) as a preliminary proof-of-concept. It is important to note that the results were checked manually to ensure the accuracy of the detection. In cases where inaccuracies were found, results from manual inspections were used. This application is a pilot test, but with further development, this approach could become a useful part of a comprehensive, multimodal roadway safety analysis.

Figure 4: Intersection Locations within Selected Census Tracts

Table 6 provides a detailed overview of changes in non-motorist facilities and safety features at selected intersections in 30 census tracts in the City of Milwaukee, WI between 2011 and 2022. These census tracts containing these intersections have a variety of median household income levels, development density, land uses, and street types. Each study intersection within these tracts has been evaluated for several key features that could impact pedestrian and cyclist safety and accessibility.

The results show that the total number of crosswalks at the 30 study intersections generally remained unchanged, indicating stability in basic pedestrian infrastructure. However, there were fluctuations in the condition of crosswalk markings. For example, Intersection 1 saw an increase in crosswalks in good condition (+1) and a decrease in those in fair condition (-1), possibly reflecting targeted maintenance efforts. Conversely, Intersection 5 experienced a decrease in crosswalks in good condition (-4) and an increase in those in poor condition (+2). Overall, based on the results, the number of crosswalks in good condition decreased during the analyzed period, highlighting the need for improved maintenance efforts.

Regarding zebra crosswalks, most intersections had no changes, indicating either stability or a lack of emphasis on this crosswalk type. For transverse crosswalks, Intersection 1 and Intersection 21 noted a significant reduction (-4), while no notable increases were reported, suggesting a potential shift away from this design. In contrast, ladder crosswalks saw increases at Intersection 1 and several others (e.g., Intersection 18), highlighting a trend towards designs that enhance visibility and safety.

Changes in the number of bicycle lanes varied, with some intersections such as Intersection 10 and Intersection 26 seeing an increase (+4), while others such as Intersection 4 and Intersection 15 saw a decrease (-1). Improvements in pavement condition were noted at several intersections (e.g., Intersection 1 and Intersection 8, both +1), which is crucial for non-motorist safety. However, some intersections experienced a decline (e.g., Intersection 5 and Intersection 13, both -2), pointing to areas needing attention. Moreover, the presence of ADA ramps increased at various intersections (e.g., Intersection 1 and Intersection 10, both +1), reflecting efforts to enhance accessibility for individuals with disabilities. However, many intersections reported no change, indicating a need for broader implementation.

In summary, the results highlight differences in the quality and distribution of pedestrian and cyclist infrastructure across different census tracts. The trends and changes related to these facilities and their quality also varied among different areas during the analysis period. Application of this type of analysis can help identify locations for tailored interventions within a jurisdication. For example, areas with lower walkability and higher transportation inequity may require targeted interventions to improve non-motorist facilities. Such areas often lack adequate infrastructure, leading to decreased safety and accessibility for pedestrians and cyclists. Addressing these disparities is crucial for promoting equitable and sustainable urban development.

Table 6. Change in Non-Motorist Facilities Between the Year 2011 and 2022

Intersection Index	Changes (“+” for increased, “-” for Decreased, “0” for no changes)											
	Number of Crosswalks	Number of Crosswalks (Marking) in Good Condition	Number of Crosswalks (Marking) in Fair Condition	Number of Crosswalks (Marking) in Poor Condition		Number of Zebra Crosswalks	Number of Transverse Crosswalks	Number of Ladder Crosswalks	Number of Bicycle		Condition of Pavement	Presence of ADA Ramps
1	0	1	-1	0		0	-4	4	1		1	1
2	0	0	0	0		0	0	0			0	0
3	0	-2	1	1		0	0	0			1	0
4	0	0	2	-2		0	0	0	-1		0	0
5	-2	-4	0	2		0	0	-2			-2	0
6	0	-4	4	0		0	0	0			0	0
7	0	-2	2	0		0	0	0			-1	0
8	0	0	0	0		0	0	0	2		1	0
9	0	0	0	0		0	0	0			0	0
10	0	4	-4	0		0	0	0	4		1	1
11	0	0	0	0		0	0	0	2		0	1
12	0	-2	2	0		0	0	0			-1	0
13	0	0	-2	2		0	0	0	3		-2	1
14	0	2	-2	0		0	-2	2	2		2	1
15	0	-2	2	0		0	0	0	-1		-1	0
16	0	2	-2	0		0	0	0	1		0	1
17	0	-4	4	0		0	0	0			-1	1
18	0	3	0	-3		0	-4	4			1	1
19	0	-2	2	0		0	0	0			1	1
20	0	-1	1	0		0	0	0			1	0
21	0	-1	1	0		0	-4	4	2		1	1
22	0	1	-1	0		0	0	0	2		1	0
23	0	-3	-1	4		0	0	0			0	1
24	0	4	0	-4		0	0	0			1	1
25	0	0	0	0		0	-2	1			0	1
26	0	4	0	-4		0	0	0	4		1	1
27	0	-1	1	0		0	0	0			1	0
28	0	-1	1	0		0	0	0			0	1
29	0	-1	3	-2		0	1	0	2		1	0
30	0	0	0	0		0	0	0			1	0

Further applications of this approach could involve comparing past and current pedestrian and bicyclist infrastructure presence and condition with safety outcomes in different neighborhoods. For example, the presence of crosswalks or other pedestrian crossing facilities could be compared with fatal and severe pedestrian crash rates (e.g., crashes per 100,000 population; crashes per trip or per walk commuter) by census tract. Facility presence and crash rates could also be compared to measures of neighborhood disadvantage, such as the Transportation Insecurity Index from the USDOT Transportation Disadvantaged Census Tracts tool, to assess the equity implications of pedestrian safety outcomes by neighborhood.

As high-quality, historic aerial images become increasingly available and vision-based large language models are refined for greater accuracy, this type of analytical approach has the potential to enhance comprehensive safety analyses for pedestrians and bicyclists. The concept illustrated in this report could help transportation agencies understand where past infrastructure investments have been made, where pedestrian facilities are improving, and where future projects are needed.

Conclusions and Recommendations

As urban environments increasingly prioritize sustainable transportation, the demand for accurate and reliable data on non-motorist activities and facilities, such as pedestrian walkways, crosswalks, and bicycle lanes, becomes crucial. This study explored recent technological advancements and methodologies for extracting non-motorist data, evaluating their ability to detect, classify, monitor changes in, and assess the condition of these facilities. Specifically, it examined the use of vision-based large language models, with a focus on ChatGPT, for detecting and evaluating non-motorist facilities from aerial images. ChatGPT, an advanced language model developed by OpenAI, is renowned for its natural language processing capabilities. With its large language model features coupled with computer vision capabilities, ChatGPT can analyze image input and generate human-like responses. By leveraging these combined capabilities, ChatGPT can interpret and generate valuable insights from complex datasets, such as aerial images.

In this study, a tool called "AI Image Analyzer" was developed to access GPT through an API. This tool mirrors the capabilities of the online version but offers additional flexibility and adaptability tailored to specific needs. To evaluate the performance of the ChatGPT in detecting and evaluating non-motorist facilities in images, a dataset containing intersection images featuring different non-motorist facilities was created. The performance of ChatGPT in various tasks was tested, and accuracy metrics were established. The result of the study showed that one of the significant advantages of using ChatGPT for extracting and evaluating non-motorist facilities from aerial images is its ability to detect and explain the location of various non-motorist facilities with reasonable accuracy. It can identify pedestrian crosswalks and describe their positions using different terminologies, such as cardinal directions or intersection corners. This flexibility in descriptive language enhances its utility in diverse urban planning contexts. Additionally, ChatGPT evaluates the condition of crosswalks and sidewalks, providing reliable assessments that can inform maintenance and improvement strategies. Moreover, its ability to explain the presence or absence of ADA ramps and street parking from aerial imagery is noteworthy. Furthermore, in assessing overall pedestrian safety and walkability for a specific area, ChatGPT considers a range of factors, including neighborhood design, green spaces, and the presence of trees, demonstrating its ability to understand the context and align with contemporary urban planning principles.

Despite its advantages, ChatGPT has a few limitations when it comes to retrieving non-motorist facilities from aerial images. The model shows inconsistency in detecting certain features, such as bicycle lanes, which it often fails to recognize without clear signage or bike images. This limitation underscores the model's reliance on explicit visual cues, which may not always be present in aerial images. Although ChatGPT can detect crosswalks, it struggles with consistently classifying them into specific types, such as traverse, ladder, or zebra crossings, which could lead to ambiguities in planning and reporting. The model's varied terminology in describing locations can also introduce confusion, necessitating additional clarification in some cases. Another critical limitation is its inability to reliably detect street lighting and pedestrian signals which can be manually recognized as their shadows casted from a satellite image. This gap in detection capabilities means that ChatGPT's evaluations might overlook crucial safety features, potentially compromising the comprehensiveness of its assessments. Therefore, while ChatGPT offers promising capabilities, its limitations highlight the need for supplementary tools and methods to achieve a thorough evaluation of non-motorist facilities.

This study utilized a limited number of satellite images to assess the performance of ChatGPT. Utilizing larger datasets can provide more comprehensive insights into the model's accuracy and applicability. Moreover, the study focused on a specific area; expanding the dataset to include images from various geographic regions and contexts can enhance the generalizability of the findings. Additionally, exploring different metrics and evaluation methods could offer more insights into ChatGPT's capabilities and limitations in retrieving non-motorist facilities from aerial images. These steps are important for advancing the reliability and effectiveness of using AI models like ChatGPT in urban planning and infrastructure assessment.

This research represents the initial phase of retrieving non-motorist facility data from aerial images using LLMs, specifically ChatGPT. It underscores the necessity of employing vision-based large language models like ChatGPT that are specifically trained on non-motorist data and with domain knowledge such as the Highway Safety Manual, Manual on Uniform Traffic Control Devices (MUTCD), Improving Intersections for Pedestrians and Bicyclists — Informational Guide, Pedestrian Safety Audit Guidelines and Prompt Lists, and rich resources at USDOT websites and others. Expanding this approach can enhance the accuracy and applicability of such models in addressing transportation safety and infrastructure assessment challenges related to non-motorists.

References

Ahmed, T., Moeinaddini, M., Almoshaogeh, M., Jamal, A., Nawaz, I., & Alharbi, F. (2021). A New Pedestrian Crossing Level of Service (PCLOS) Method for Promoting Safe Pedestrian Crossing in Urban Areas. *International Journal of Environmental Research and Public Health*, *18*(16), 8813. https://doi.org/10.3390/ijerph18168813

Anciaes, P. R., & Jones, P. (2018). Estimating preferences for different types of pedestrian crossing facilities. *Transportation Research Part F: Traffic Psychology and Behaviour*, *52*, 222–237. https://doi.org/10.1016/j.trf.2017.11.025

Arora, S., Narayan, A., Chen, M. F., Orr, L., Guha, N., Bhatia, K., Chami, I., Sala, F., & Ré, C. (2022). *Ask Me Anything: A simple strategy for prompting language models*. 1–63.

Baučić, M., Gilić, F., Bačić, S., & Duplančić-Leder, T. (2020). OPEN GEOSPATIAL DATA FOR URBAN GREEN AREAS. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *VI-4/W2-2020*, 17–24. https://doi.org/10.5194/isprs-annals-VI-4-W2-2020-17-2020

Bian, Y., Liang, K., Zhao, X., Li, H., & Yang, L. (2020). Evaluating the effectiveness of new-designed crosswalk markings at intersections in China considering vehicle-pedestrian interaction. *Accident Analysis & Prevention*, *139*, 105498. https://doi.org/10.1016/j.aap.2020.105498

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

Carranza-García, M., Torres-Mateo, J., Lara-Benítez, P., & García-Gutiérrez, J. (2021). On the Performance of One-Stage and Two-Stage Object Detectors in Autonomous Vehicles Using Camera Data. *Remote Sensing*, *13*(1). https://doi.org/10.3390/rs13010089

Chen, Z., Luo, R., Li, J., Du, J., & Wang, C. (2021). U-Net Based Road Area Guidance for Crosswalks Detection from Remote Sensing Images. *Canadian Journal of Remote Sensing*, *47*(1), 83–99. https://doi.org/10.1080/07038992.2021.1894915

Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., Barham, P., Chung, H. W., Sutton, C., & Gehrmann, S. (2023). Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research*, *24*(240), 1–113.

Das, S., Tamakloe, R., Zubaidi, H., Obaid, I., & Alnedawi, A. (2021). Fatal pedestrian crashes at intersections: Trend mining using association rules. *Accident Analysis & Prevention*, *160*, 106306. https://doi.org/10.1016/j.aap.2021.106306

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805). arXiv. http://arxiv.org/abs/1810.04805

Fan, Y., Sun, Z., & Zhao, G. (2020). A Coarse-to-Fine Framework for Multiple Pedestrian Crossing Detection. *Sensors*, *20*(15). https://doi.org/10.3390/s20154144

Feldman, M., Manzi, J. G., & Mitman, M. F. (2010). Empirical Bayesian Evaluation of Safety Effects of High-Visibility School (Yellow) Crosswalks in San Francisco, California. *Transportation Research Record*, *2198*(1), 8–14. https://doi.org/10.3141/2198-02

Francke, A. (2022). Cycling during and after the COVID-19 pandemic. *Advances in Transport Policy and Planning*, *10*, 265–290. https://doi.org/10.1016/bs.atpp.2022.04.011

Gaw, L., Chen, S., Chow, Y. S., Lee, K., & Biljecki, F. (2022). COMPARING STREET VIEW IMAGERY AND AERIAL PERSPECTIVES IN THE BUILT ENVIRONMENT. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *X-4/W3-2022*, 49–56. https://doi.org/10.5194/isprs-annals-X-4-W3-2022-49-2022

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 580–587. https://doi.org/10.1109/CVPR.2014.81

Haddad, A. J., Mondal, A., Bhat, C. R., Zhang, A., Liao, M. C., Macias, L. J., Lee, M. K., & Watkins, S. C. (2023). Pedestrian crash frequency: Unpacking the effects of contributing factors and racial disparities. *Accident Analysis & Prevention*, *182*, 106954. https://doi.org/10.1016/j.aap.2023.106954

Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., Casas, D. de L., Hendricks, L. A., Welbl, J., Clark, A., Hennigan, T., Noland, E., Millican, K., Driessche, G. van den, Damoc, B., Guy, A., Osindero, S., Simonyan, K., Elsen, E., … Sifre, L. (2022). *Training Compute-Optimal Large Language Models* (arXiv:2203.15556). arXiv. http://arxiv.org/abs/2203.15556

Hossain, A., Sun, X., Zafri, N. M., & Codjoe, J. (2023). Investigating pedestrian crash patterns at high-speed intersection and road segments: Findings from the unsupervised learning algorithm. *International Journal of Transportation Science and Technology*. https://doi.org/10.1016/j.ijtst.2023.04.007

Huang, P., Wang, S., Chen, J., Li, W., & Peng, X. (2023). Lightweight Model for Pavement Defect Detection Based on Improved YOLOv7. *Sensors*, *23*(16), Article 16. https://doi.org/10.3390/s23167112

Ju, C., Han, T., Zheng, K., Zhang, Y., & Xie, W. (2022). Prompting Visual-Language Models for Efficient Video Understanding. In S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, & T. Hassner (Eds.), *Computer Vision – ECCV 2022* (Vol. 13695, pp. 105–124). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-19833-5_7

Jung, J., Qin, L., Welleck, S., Brahman, F., Bhagavatula, C., Le Bras, R., & Choi, Y. (2022). Maieutic Prompting: Logically Consistent Reasoning with Recursive Explanations. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022*, 1266–1279. https://doi.org/10.18653/v1/2022.emnlp-main.82

Karaer, A., Kaczmarek, W., Mank, E., Ghorbanzadeh, M., Koloushani, M., Dulebenets, M. A., Moses, R., Sando, T., & Ozguven, E. E. (2023). Traffic Data on-the-Fly: Developing a Statewide Crosswalk Inventory Using Artificial Intelligence and Aerial Images (AI2) for Pedestrian Safety Policy Improvements in Florida. *Data Science for Transportation*, *5*(2), 7. https://doi.org/10.1007/s42421-023-00070-1

Karimi, H. A., & Kasemsuppakorn, P. (2013). Pedestrian network map generation approaches and recommendation. *International Journal of Geographical Information Science*, *27*(5), 947–962. https://doi.org/10.1080/13658816.2012.730148

Kaya, Ö., Çodur, M. Y., & Mustafaraj, E. (2023). Automatic Detection of Pedestrian Crosswalk with Faster R-CNN and YOLOv7. *Buildings*, *13*(4). https://doi.org/10.3390/buildings13041070

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 25). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

Lee, C., & Abdel-Aty, M. (2005). Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. *Accident Analysis & Prevention*, *37*(4), 775–786. https://doi.org/10.1016/j.aap.2005.03.019

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *ACM Computing Surveys*, *55*(9). https://doi.org/10.1145/3560815

Mitman, M. F., Ragland, D. R., & Zegeer, C. V. (2008). Marked-Crosswalk Dilemma: Uncovering Some Missing Links in a 35-Year Debate. *Transportation Research Record*, *2073*(1), 86–93. https://doi.org/10.3141/2073-10

Moran, M. E. (2022). Where the Crosswalk Ends: Mapping Crosswalk Coverage via Satellite Imagery in San Francisco. *Environment and Planning B: Urban Analytics and City Science*, *49*(8), 2250–2266. https://doi.org/10.1177/23998083221081530

National Center for Statistics and Analysis. (2023). *Pedestrians: 2021 data (Traffic Safety Facts. Report No. DOT HS 813 458).* National Highway Traffic Safety Administration.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, *1*(8), 9.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, *21*(140), 1–67.

Reynolds, L., & McDonell, K. (2021). Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3411763.3451760

Roll, J., & McNeil, N. (2022). Race and income disparities in pedestrian injuries: Factors influencing pedestrian safety inequity. *Transportation Research Part D: Transport and Environment*, *107*, 103294. https://doi.org/10.1016/j.trd.2022.103294

Sharma, V., Gupta, M., Kumar, A., & Mishra, D. (2021). Video Processing Using Deep Learning Techniques: A Systematic Literature Review. *IEEE Access*, *9*, 139489–139507. https://doi.org/10.1109/ACCESS.2021.3118541

Shin, T., Razeghi, Y., Logan, R. L., Wallace, E., & Singh, S. (2020). AUTOPROMPT: Eliciting knowledge from language models with automatically generated prompts. *EMNLP 2020 - 2020 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, 4222–4235. https://doi.org/10.18653/v1/2020.emnlp-main.346

Shoaib, M. R., Emara, H. M., & Zhao, J. (2023). A survey on the applications of frontier ai, foundation models, and large language models to intelligent transportation systems. *2023 International Conference on Computer and Applications (ICCA)*, 1–7. https://ieeexplore.ieee.org/abstract/document/10401518/

Smith, S., Patwary, M., Norick, B., LeGresley, P., Rajbhandari, S., Casper, J., Liu, Z., Prabhumoye, S., Zerveas, G., Korthikanti, V., Zhang, E., Child, R., Aminabadi, R. Y., Bernauer, J., Song, X., Shoeybi, M., He, Y., Houston, M., Tiwary, S., & Catanzaro, B. (2022). *Using DeepSpeed and Megatron to Train Megatron-Turing NLG 530B, A Large-Scale Generative Language Model* (arXiv:2201.11990). arXiv. http://arxiv.org/abs/2201.11990

Taori, R., Gulrajani, I., Zhang, T., Dubois, Y., Li, X., Guestrin, C., Liang, P., & Hashimoto, T. B. (2023). *Stanford alpaca: An instruction-following llama model*.

Tian, D., Han, Y., Wang, B., Guan, T., & Wei, W. (2021). A Review of Intelligent Driving Pedestrian Detection Based on Deep Learning. *Computational Intelligence and Neuroscience*, *2021*, 5410049. https://doi.org/10.1155/2021/5410049

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA: Open and Efficient Foundation Language Models* (arXiv:2302.13971). arXiv. http://arxiv.org/abs/2302.13971

Tümen, V., & Ergen, B. (2020). Intersections and crosswalk detection using deep learning and image processing techniques. *Physica A: Statistical Mechanics and Its Applications*, *543*, 123510. https://doi.org/10.1016/j.physa.2019.123510

Twaddell, H., Rose, E., Broach, J., Dill, J., Clifton, K. J., Lust, C., Voros, K., Louch, H., David, E., ICF, Portland State University, & Alta Planning + Design. (2018). *Guidebook for Measuring Multimodal Network Connectivity* (FHWA-HEP-18-032). https://rosap.ntl.bts.gov/view/dot/50786

van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: five priorities for research. *Nature*, *614*(7947), 224–226. https://doi.org/10.1038/d41586-023-00288-7

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*. https://proceedings.neurips.cc/paper/7181-attention-is-all

Verma, R., & Ukkusuri, S. V. (2023). Crosswalk Detection from Satellite Imagery for Pedestrian Network Completion. *Transportation Research Record*, 03611981231210545. https://doi.org/10.1177/03611981231210545

Vicuna: An Open-Source Chatbot Impressing GPT-4 with 90% ChatGPT Quality | LMSYS Org*. (n.d.). Retrieved June 26, 2024, from https://lmsys.org/blog/2023-03-30-vicuna

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022). *Emergent Abilities of Large Language Models*. 1–30.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E. H., Le, Q. V., & Zhou, D. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. *Advances in Neural Information Processing Systems*, *35*(NeurIPS), 1–14.

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT*.

World Health Organization. (2015). *Global status report on road safety 2015*. World Health Organization.

Wu, X., Sahoo, D., & Hoi, S. C. H. (2020). Recent advances in deep learning for object detection. *Neurocomputing*, *396*, 39–64. https://doi.org/10.1016/j.neucom.2020.01.085

Xin, M., & Wang, Y. (2019). Research on image classification model based on deep convolution neural network. *EURASIP Journal on Image and Video Processing*, *2019*(1), 40. https://doi.org/10.1186/s13640-019-0417-8

Zegeer, C. V., Stewart, J. R., Huang, H. H., Lagerwey, P. A., Feaganes, J. R., & Campbell, B. (2005). *Safety effects of marked versus unmarked crosswalks at uncontrolled locations final report and recommended guidelines*. United States. Federal Highway Administration. Office of Safety Research and ….

Zhang, D., Zhang, W., Li, F., Liang, K., & Yang, Y. (2023). PNANet: Probabilistic Two-Stage Detector Using Pyramid Non-Local Attention. *Sensors (Basel, Switzerland)*, *23*(10), 4938. https://doi.org/10.3390/s23104938

Zhang, S., Roller, S., Goyal, N., Artetxe, M., Chen, M., Chen, S., Dewan, C., Diab, M., Li, X., Lin, X. V., Mihaylov, T., Ott, M., Shleifer, S., Shuster, K., Simig, D., Koura, P. S., Sridhar, A., Wang, T., & Zettlemoyer, L. (2022). *OPT: Open Pre-trained Transformer Language Models* (arXiv:2205.01068). arXiv. http://arxiv.org/abs/2205.01068

Zhang, Y., Zhang, C., & Luttrell, J. (2021). *An Automated System for Pedestrian Facility Data Collection from Aerial Images* (NCHRP IDEA Project 209.).

Zheng, O., Abdel-Aty, M., Wang, D., Wang, Z., & Ding, S. (2023). ChatGPT is on the horizon: Could a large language model be suitable for intelligent traffic safety research and applications? *arXiv Preprint arXiv:2303.05382*.

Zhou, Y., Muresanu, A. I., Han, Z., Paster, K., Pitis, S., Chan, H., & Ba, J. (2022). *Large Language Models Are Human-Level Prompt Engineers*.