Mitigating out-of-distribution Challenges in Connected Vehicle-Infrastructure-Pedestrian Systems for Promoting Pedestrian Safety at Intersections

A Technical Report Submitted to the Rural Safe Efficient Advanced Transportation (R-SEAT) Center and United States Department of Transportation

FINAL REPORT

Principal Investigator:
Susu Xu, Ph.D.
Affiliated Assistant Professor
Department of Civil Engineering
Stony Brook University
Assistant Professor
Department of Civil and Systems Engineering
Johns Hopkins University
E-mail: susuxu@jhu.edu

Research Assistant:
Chenguang Wang, M. Sc.
Ph.D. Student
Department of Civil Engineering
Stony Brook University
1208 Old Computer Science Building
E-mail: Chenguang.wang@stonybrook.edu

February 2025

DISCLAIMER

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle Mitigating Out-of-distribution Challenges in Connected Vehicle-Infrastructure-Pedestrian Systems for Promoting Pedestrian Safety at Intersections	5. Report Date 02/24/2025
	6. Performing Organization Code 59-0977035
7. Author(s) Susu Xu https://orcid.org/0000-0001-7170-648X Chenguang Wang https://orcid.org/0009-0006-6445-6703	8. Performing Organization Report No.
9. Performing Organization Name and Address Stony Brook University	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. 69A3552348321
12. Sponsoring Agency Name and Address Rural Safe Efficient Advanced Transportation (R-SEAT) Center 2525 Pottsdamer Street Tallahassee, FL 32310	13. Type of Report and Period Covered Final Report Period Covered: 02/22/2023 – 12/31/2024
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Pedestrian trajectory prediction constitutes a critical component of motion planning in autonomous vehicles (AVs), facilitating safe and efficient navigation. While recent advancements in deep learning have enabled near real-time prediction capabilities, ethical concerns have emerged within the artificial intelligence (AI) community regarding algorithmic bias and fairness. In particular, the majority of publicly available pedestrian trajectory datasets predominantly represent majority populations, raising concerns about the generalizability of trained models to diverse pedestrian groups, especially vulnerable populations such as individuals with disabilities, older adults, and children. Such biases in prediction models may disproportionately impact these groups, potentially increasing their risk of involvement in vehicular accidents. In this study, we evaluate two state-of-the-art pedestrian trajectory prediction models for potential biases associated with age and gender, utilizing three distinct datasets. We introduce novel evaluation metrics tailored to assess differential model performance across demographic subgroups. Our empirical findings reveal that both models demonstrate reduced predictive accuracy for children and older adults relative to adults, while exhibiting comparable performance across male and female pedestrians. We further analyze potential sources of observed biases and delineate key limitations of our study. Future work will focus on expanding model evaluations, refining bias assessment methodologies, disentangling dataset-related biases from model-induced biases, and developing strategies for mitigating algorithmic bias.

17. Key Words Autonomous driving, Pedestrian trajectory prediction, AI model bias	18. Distribution Statement No restrictions

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

ACKNOWLEDGEMENTS

This project was sponsored by the Rural Safe Efficient Advanced Transportation (R-SEAT) Center and United States Department of Transportation. The Principal Investigators would like to thank the representatives of the R-SEAT Center for their valuable feedback throughout the project activities.

EXECUTIVE SUMMARY

This project introduces a novel fairness-aware, multi-modal pedestrian safety system that leverages connected vehicle-infrastructure-pedestrian (VIP) networks to enhance the safety and quality of experience for all pedestrians. The primary objective is to uncover, characterize, and mitigate algorithmic biases embedded in artificial intelligence (AI) models utilized in VIP systems through advanced machine learning and representation learning approaches, using real-world video data and simulation platforms.

The research is structured into three core thrusts:

1. Uncovering and Characterizing Pedestrian Intent Variations and Out-of-distribution (OOD) Challenges:

The first thrust investigates variations in pedestrian intent across demographic groups—defined by age, gender, and disability—by analyzing variations in movement trajectories, speeds, and gait patterns using three annotated video datasets: JAAD, PIE, and TITAN. Metrics such as hypothesis testing-based scores and Wasserstein distances are used to quantify these variations. The study also evaluates the OOD problems present in both deterministic and probabilistic trajectory prediction models by testing model performance across demographic groups and quantifying fairness metrics.

2. Developing Algorithms to Mitigate OOD Impacts:

This thrust addresses OOD problems arising from data scarcity and distribution shifts that disproportionately affect vulnerable pedestrian populations. The team will develop fairness-aware, multi-modal learning frameworks that use distributional divergence metrics (e.g., Wasserstein distance, KL divergence) to detect and characterize data imbalances. A multi-view graph-based model informed by physical interdependencies between system entities (vehicles, infrastructure, pedestrians) will be constructed. Advanced methods—including Bayesian inference, deep graph neural networks, adversarial on-manifold data augmentation, and fairness-aware loss functions—will be employed to jointly optimize pedestrian detection and trajectory prediction while minimizing group-level discrepancies in model performance.

3. Building a Simulation Platform for Intersection Evaluation:

To validate the proposed methodologies, the project will develop a simulation platform based on CARLA, an open-source virtual environment for autonomous driving research. In Phase 1, real-world data will be collected through mobile and vehicle-based sensors to inform model development. In Phase 2, this data will be used to simulate various intersection scenarios, integrating perception, control, and pedestrian modules. The system's performance will be assessed under different conditions with and without the proposed fairness-aware framework, allowing for a comprehensive evaluation of safety and OOD mitigation outcomes.

In summary, this project delivers a comprehensive framework for understanding and mitigating OOD challenges in AI-driven pedestrian safety systems. By integrating multi-modal sensing, fairness-aware learning, and large-scale simulation, the proposed work aims to promote reliability and safety for all pedestrian populations within next-generation intelligent transportation systems.

2 Table of Contents

1. INTRODUCTION

1.1. Pedestrian Safety

Pedestrian safety continues to pose a major challenge in today's transportation systems. In the United States, fatal pedestrian collisions have surged by nearly 50% over the past decade [27]. However, experts in transportation emphasize that these incidents should not be viewed merely as "accidents," but rather as outcomes rooted in deep-seated systemic transportation safety concerns within our society.

Autonomous vehicles (AVs) are poised to fundamentally transform pedestrian-vehicle interactions through their sophisticated control capabilities. Recent advancements have led to the development of various collision prevention systems that identify pedestrians and forecast their movements using sensor data from on-board cameras and machine learning algorithms [24, 36]. These innovations allow AVs to detect pedestrians in real-time and respond to emerging hazards. For instance, Waymo reports that its 5th-generation Waymo Driver can identify pedestrians and traffic signs at distances up to 500 meters [16]. Such pedestrian perception systems rely heavily on rich sensory input and high-capacity AI models capable of learning complex behavioral cues.

Nevertheless, out-of-distribution problems remain pressing concerns in contemporary AI systems [2, 22, 31, 35]. With the proliferation of Internet of Things (IoT) technologies, a wide array of cameras and smart sensors are now embedded in vehicles to monitor environmental surroundings [7, 38, 39]. Yet, due to mobility constraints and limited data availability, vulnerable pedestrian groups—particularly children and older adults—are underrepresented in training datasets. These groups often exhibit behavioral patterns that differ significantly from the general pedestrian population, contributing to discrepancies in detection accuracy and safety outcomes.

1.2. Project Objectives

This project seeks to uncover, characterize, and mitigate the out-of-distribution challenges of AI models in connected vehicle-infrastructure-pedestrian systems by leveraging advanced statistical machine learning and representation learning techniques and real-world video data. Specifically, we will focus on three tasks:

(1) To uncover and characterize pedestrian intent by leveraging multiple sensing modalities and heterogeneous data available from mobile devices carried by pedestrians and sensors mounted on vehicles and infrastructures

(2) To design a pipeline for evaluating system fairness in pedestrian trajectory predictions and provide insights into the discriminatory system behaviors.

(3) To develop novel fairness-aware multi-modal machine learning algorithms, in supervised and semi-supervised ways, to achieve fair and accurate support for all groups of pedestrians

1.3. Project Relevance to the REAT Themes and USDOT Strategic Plan

As a REAT-funded study, this project introduce a novel set of metrics and framework to enhance the quality of experience of pedestrians by leveraging connected vehicle-infrastructure-pedestrian (VIP). The project well supports USDOT priorities and RD&T strategic goals, especially on the dimensions of safety. The methodological development is transformative, leading to impactful use of datasets maintained by USDOT and public datasets contributed by the research community.

1.4 Organization of the Report

This report is organized to guide the readers through the project's major activities. In the next section, the report presents the causal modeling and extracted data evidence for evaluating the mobility of seniors living in rural areas. Following that, Section 3 summarizes the data analysis for determining the vulnerability of elderly drivers. Upon the insights learnt from these two sections, Section 4 further presents a proposed method in addressing an impressing challenge in creating situational awareness in complex driving scenes. In the end, Section 5 summarizes major findings and outputs from this project.

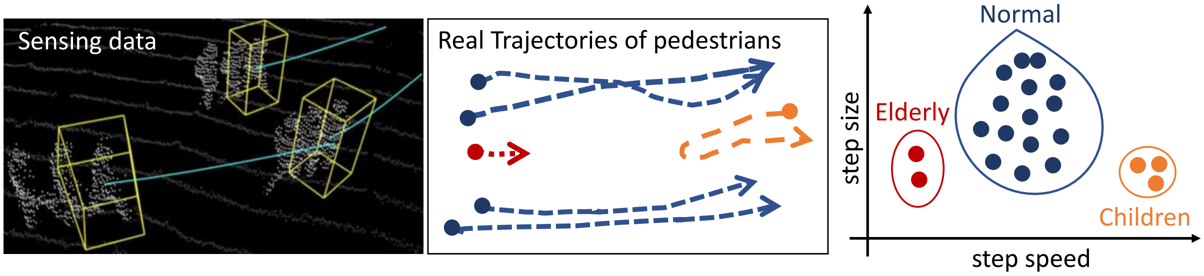

Figure 1: Concepts of out-of-distribution behavior patterns of vulnerable pedestrians. (a) Majority of pedestrian data are from normal pedestrians. (b) The vulnerable pedestrians such as elder/disabled (red) and children (orange) often have distinct crossing behavior patterns resulting in different trajectories. (c) Compared to data distribution of majority normal pedestrians, the data of vulnerable groups are often ``minor mode" or ``out-of-distribution".

2. Background

Pedestrians have historically been marginalized in the design of transportation infrastructure, which has primarily prioritized motor vehicle traffic. As urbanization accelerates across the United States [6] and public health campaigns increasingly promote active lifestyles [7, 8], pedestrian activity has surged significantly—a trend projected to persist in the coming years [9]. Alarming statistics from the National Highway Traffic Safety Administration (NHTSA) indicate that 2018 marked the highest number of pedestrian fatalities in the U.S. since 1990 [10, 11]. Vulnerable populations—such as the elderly, people with disabilities, and children—are particularly at risk, with research showing that pedestrians using wheelchairs are 36% more likely to be fatally struck by a vehicle than those who are not [12]. Additionally, inadequate pedestrian support at intersections often leads to jaywalking, a well-documented contributor to urban pedestrian deaths [9, 11]. These concerns highlight the urgent need for transportation systems that prioritize pedestrian-centric services, especially for children, seniors, and individuals with disabilities.

The development of effective pedestrian safety systems must tackle two primary challenges to ensure fairness across all road users: (1) algorithmic out-of-distribution arising from the limited and behaviorally distinct data associated with vulnerable pedestrians, and (2) the difficulty of balancing the needs of pedestrians with those of vehicle occupants. Recently, several collision avoidance systems—installed either on vehicles or infrastructure—have emerged to detect pedestrians and anticipate their movements using sensory inputs such as LiDAR and cameras, combined with machine learning algorithms [13, 14]. These systems depend on extensive training data and high-capacity models to capture the complexities of pedestrian behavior. However, the data representing vulnerable pedestrians is often scarce and characterized by unique behavioral patterns compared to more commonly represented groups [15, 16]. For instance, unaccompanied children and elderly pedestrians exhibit markedly different walking speeds and path trajectories than young adults, and are significantly underrepresented in training datasets [16]. This scarcity and distributional variation cause the data from certain groups as scarce “modes” or even “out-of-distribution” cases within machine learning models, leading to elevated error rates during detection and prediction for these groups. These inaccuracies can result in unsafe outcomes—for example, an intelligent traffic light system optimized for the average pedestrian may fail to accommodate the needs of underrepresented vulnerable pedestrians, thereby posing serious safety risks.

However, out-of-distribution are critical problems that exist in many AI systems today [10,11,12,13]. With the development of the Internet of Things, various cameras and smart devices are deployed on vehicles to capture the data from the surrounding environment. However, due to the constrained mobilities of some types of road users, data from these less dominated pedestrians, such as children, elders, and disabled, is often limited. These pedestrians have distinct distribution patterns compared to other pedestrian groups [14]. For example, children are more likely to exhibit unpredictable behaviors [15], and elderly pedestrians on average walk slower than the general population [16], as shown conceptually in Figure 1. Additionally, men tend to display more risky and impulsive behaviors than women [17]. The data scarcity and distinct distributions of these pedestrian groups will make their data minor ``mode" or even ``out-of-distribution" compared to the huge amount of training data from other pedestrian groups. This will lead to larger prediction errors during the testing stage. Error-prone detection and trajectory prediction of vulnerable pedestrians may cause discriminatory decision-making against these groups, compromising their safety. Previous works exploring distribution shift problems in AV perception primarily focused on pedestrian detection. Hirota et al. [18] and Kogure et al. [19] found that state-of-the-art models had lower detection rates for children compared adults.

To understand and address these out-of-distribution challenges, we propose a novel multi-modal pedestrian safety system to enhance the quality of experience of pedestrians, by leveraging connected vehicle-infrastructure-pedestrian (VIP).

3 ALGORITHMIC BIAS IN AUTONOMOUS TRAJECTORY PREDICTION

Predicting future trajectories equips autonomous vehicles (AVs) with crucial information required for navigating safely through dynamic and interactive environments, thus significantly reducing the risk of collisions or near-miss incidents. Despite its importance, trajectory prediction remains inherently challenging due to the myriad of variables influencing pedestrian movements in real-time, such as social interactions, environmental conditions, and individual intentions. Initially, researchers addressed pedestrian trajectory prediction by developing rule-based or physics-inspired models that simulated pedestrian behavior using straightforward principles and mechanisms, such as repulsive forces to avoid collisions or interactions modeled as physical interactions between particles [1, 2]. However, recent advances in deep learning technologies have shifted the research paradigm towards data-driven models. Typically, trajectory prediction data is derived from video footage captured by sensors, either from vehicle-mounted cameras (on-board view) or aerial views (bird's-eye view). In processing these videos, researchers manually or algorithmically annotate each frame, marking pedestrians with bounding boxes. Mathematically, the trajectory data at any given time t comprises past τ observations expressed as:

$$X_t = \left[x_{\{t-\tau+1\}}, x_{\{t-\tau+2\}}, \dots, x_{\{t\}}\right],$$

where X_t represents the bounding box coordinates of a pedestrian at time t. The primary objective of trajectory prediction algorithms is to accurately forecast future pedestrian locations, represented as bounding box coordinates:

$$Y_t = \left[y_{\{t+1\}}, y_{\{t+2\}}, \dots, y_{\{t+\delta\}}\right],$$

for a future interval of δ frames.

3.1 Automated Trajectory Prediction for Autonomous Driving

We assess two state-of-the-art trajectory prediction models known for their strong performance on existing benchmarks: BiTraP [7] and SGNet [8]. Both models have multiple variants designed for different prediction scenarios. BiTraP (Bidirectional Trajectory Prediction) is a goal-conditioned, multimodal prediction framework leveraging conditional variational autoencoders (CVAEs). Its primary variants include BiTraP-D (deterministic), BiTraP-NP (non-parametric multimodal), and BiTraP-GMM (Gaussian Mixture Model multimodal). In contrast, SGNet (Stepwise Goal Network) introduces a different prediction approach by estimating future goals incrementally rather than predicting a single, long-term destination. SGNet variants include a deterministic version (SGNet) and a multimodal version enhanced with conditional variational autoencoder frameworks. Both models were selected for their demonstrated effectiveness on the JAAD and PIE datasets, and because their implementations are publicly accessible, facilitating transparency and reproducibility.

3.1.1 Deterministic Models

Deterministic trajectory prediction models produce a single, definitive prediction of a pedestrian's future trajectory based solely on previously observed behavior. These models assume that past observations provide sufficient context to reliably predict future actions without accounting explicitly for uncertainty or variability. Typically, deterministic models rely on recurrent neural networks (RNNs) or transformer architectures that extract temporal dependencies and patterns from past trajectory data. In this work, we specifically evaluate deterministic variants, BiTraP-D and SGNet. These models integrate various mechanisms such as goal-conditioning or incremental goal prediction to enhance the accuracy and relevance of their predictions. In future investigations, we will also evaluate other deterministic frameworks like PIEtraj_{traj} [5], aiming to broaden our understanding of performance dynamics across different prediction strategies.

3.1.2 Multi-modal Models

Multimodal trajectory prediction models address the inherent uncertainty and complexity of pedestrian behavior by generating multiple plausible future trajectories. Unlike deterministic approaches, multimodal models explicitly acknowledge the stochastic nature of pedestrian actions and environmental interactions. These models commonly employ probabilistic techniques, such as conditional variational autoencoders (CVAEs) and Gaussian Mixture Models (GMMs), to capture the distribution of potential future paths. The multimodal approach offers several predictive hypotheses, enabling AVs to anticipate a range of possible pedestrian behaviors and thus enhancing their decision-making robustness. Our current analysis specifically explores the performance of BiTraP-NP, a non-parametric multimodal variant. Future work will also incorporate comprehensive evaluations of BiTraP-GMM and $SGNetCVAE_{\{CVAE\}}$, following established practices that select the best-performing prediction from a set of multiple sampled trajectories, commonly referred to as the "best-of-20" evaluation strategy [9,10,11].

3.2 Out-of-distribution Challenges and Algorithmic bias

Algorithmic bias has become a major area of research interest within the AI community, particularly due to increasing societal concerns about biases embedded in algorithmic decisions. Despite numerous efforts, a universally applicable definition of fairness remains elusive because fairness often depends on context-specific ethical considerations. Among various definitions, statistical parity (also referred to as demographic parity) emerges prominently for applications like pedestrian trajectory prediction. Statistical parity ensures fairness by requiring equal prediction probabilities across different demographic groups. Formally, statistical parity is mathematically defined as:

$$P\left(\hat{Y} \mid A = a \right) = P\left(\hat{Y} \mid A = b \right)$$

where $\hat{Y}$ symbolizes the predicted outcome, while aa and bb denote distinct demographic groups, such as different genders or age groups. In practical terms, statistical parity in our context implies that the prediction accuracy of trajectories does not significantly differ across various demographic groups, thus ensuring safety outcomes for pedestrians regardless of their demographic characteristics [3].

3.3 Common Datasets

In this study, we employ three distinct and carefully selected datasets, each collected via a single on-board camera and annotated specifically for pedestrian trajectory prediction: Joint Attention in Autonomous Driving (JAAD) [4], Pedestrian Intention Estimation (PIE) [5], and Trajectory Inference using Targeted Action priors Network (TITAN) [6].

The JAAD dataset primarily captures urban pedestrian interactions and behaviors in Kremenchuk, Ukraine, with supplementary data from diverse locations across Canada, Germany, the United States, and other Ukrainian cities. JAAD annotations occur at a frequency of 30 frames per second (Hz), encompassing 2,580 pedestrians overall, although detailed demographic information (age and gender) is available for approximately 648 pedestrians, about 25% of the dataset. The PIE dataset is more geographically homogeneous, recorded exclusively in Toronto, Canada, and similarly annotated at 30 Hz. It comprises demographic labels (both age and gender) for all its 1,835 pedestrians, allowing for detailed fairness analyses. Finally, the TITAN dataset, recorded entirely in Tokyo, Japan, captures pedestrian behavior at a lower annotation frequency of 10 Hz. It includes the largest number of pedestrians among our selected datasets—8,588 individuals—all annotated with age information, although notably lacking gender labels. These datasets were specifically chosen due to their rich demographic annotations, unique geographical diversity, and the availability of comprehensive metadata, enabling robust exploration of OOD problems in trajectory prediction.

Table 1: Age demographic breakdown of JAAD, PIE, and TITAN datasets. We define children to be ages 0-14, adults to be ages 15-64, and elderly to be ages 65 and above. The JAAD dataset was filmed in 5 different cities across Ukraine, Canada, Germany, and the US, but 80\% of the clips were filmed in Kremenchuk, Ukraine. We could not find age demographic data for Kremenchuk specifically, so we report the expected age demographics of Ukraine as a country.

Dataset	Statistic	Children	Adults	Elderly
JAAD	# of pedestrians	47	509	92
	% of dataset	7.3%	78.5%	14.2%
	% expected [8]	15.1%	69.3%	15.6%
PIE	# of pedestrians	17	1640	185
	% of dataset	0.9%	89.0%	10.0%
	% expected [11]	14.2%	68.1%	17.6%
TITAN	# of pedestrians	116	7872	506
	% of dataset	1.4%	91.7%	5.9%
	% expected [6]	11.5%	65.7%	22.8%

4 OUT-OF-DISTRIBUTION DETECTION, MEASUREMENT, AND MITIGATION

4.1 Definition of Pedestrian Tracks

In the context of pedestrian trajectory prediction, we define a "track" as a sequential record of the bounding box coordinates that encapsulate the position and movement of a pedestrian over a specified interval. Each track is standardized to span two seconds, segmented into two critical phases: an initial observation period of 0.5 seconds, during which pedestrian behavior is monitored and recorded, and a subsequent prediction interval of 1.5 seconds, where the trajectory prediction model forecasts future movements. This division allows models sufficient observational data to make informed predictions, reflecting typical human response times and vehicle reaction capabilities in real-world autonomous driving scenarios.

To assess the predictive accuracy of trajectory models, we adopt the mean squared error (MSE), a standard evaluation metric widely utilized in trajectory prediction research. MSE quantifies accuracy by computing the average squared differences between predicted and actual bounding box coordinates at specific time frames. Formally, the MSE at frame t is defined as:

$$MSE_t = \frac{1}{N} \sum_{\{i=1\}}^{\{N\}} \frac{1}{N} \left(y_{\{i,t\}} - \hat{y}_{\{i,t\}} \right)^2$$

where NN is the number of pedestrian samples, $y_{\{i,t\}}$ represents the actual bounding box coordinates, and $\hat{y}_{\{i,t\}}$ represents the predicted coordinates at time t. Although straightforward, MSE can disproportionately reflect the impact of outliers, potentially skewing the assessment of a model's predictive capabilities.

4.2 Algorithmic Bias Evaluation Metrics

Fairness evaluation metrics quantify how equally or unequally prediction models perform across various demographic groups. In our study, we employ three comprehensive fairness metrics: mean MSE, Mann-Whitney U Test, and Wasserstein distance.

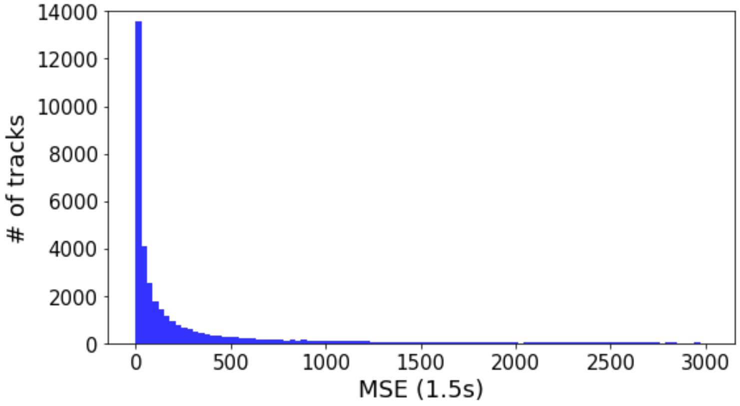

Figure 2: MSE (1.5s) error distribution of BiTraP-D performance on the PIE dataset. The skewed right trend is consistent for all models across all datasets. Note that some MSE (1.5s) values here go up to nearly 100,000, but we only plot up to 3000 in this figure.

4.2.1 Mean MSE:

Traditionally, pedestrian trajectory prediction performance on datasets such as JAAD and PIE has been assessed using three variations of mean MSE: 1) bounding box MSE averaged over intervals (0.5 s, 1.0 s, 1.5 s), 2) bounding box center mean squared error (CMSEC_{MSE}), and 3) bounding box center final mean squared error (CFMSECF_{MSE}). While widely used, these metrics' vulnerability to outliers suggests the necessity for supplementary statistical measures to ensure an accurate assessment of model performance across demographic groups.

4.2.2 Mann-Whitney U Test:

The Mann-Whitney U Test is a nonparametric test that determines whether two independent samples derive from the same population. Unlike parametric tests such as the t-test, the Mann-Whitney U Test does not require the assumption of normally distributed data, making it particularly valuable in analyzing datasets with skewed or unknown distributions, as often encountered in real-world scenarios like pedestrian trajectory prediction errors. The test ranks all observations from both groups together and calculates the sum of ranks for each group. These rank sums are then used to compute the U statistic, which measures the degree of separation between the two samples. A significant U statistic suggests that one group's observations tend to be larger or smaller than those of the other group, indicating a systematic difference in the distributions being compared.

As stated before, the error distributions in our study are not normally distributed, so we cannot use a parametric test like the t-test. We conduct two different onesided Mann-Whitney U Tests with the following null and alternative hypotheses:

- **Children vs. Adults:**
    - $\boldsymbol{H_0}$: Performance for children equals adults.
    - $\boldsymbol{H_1}$: Performance for children is worse than adults.
- **Elderly vs. Adults:**
    - $\boldsymbol{H_0}$: Performance for elderly equals adults.
    - $\boldsymbol{H_1}$: Performance for elderly is worse than adults.

In the context of fairness evaluation in pedestrian trajectory prediction, the Mann-Whitney U Test is used to examine potential performance variations across demographic groups, such as children versus adults, elderly versus adults, and men versus women. By clearly defining null and alternative hypotheses—for example, testing whether the model performance on children is significantly worse than on adults—the Mann-Whitney U Test provides rigorous statistical evidence regarding the presence of demographic biases. This approach enables researchers to identify and quantify fairness issues systematically, guiding the development trajectory prediction models and fostering deeper insights into the demographic-specific challenges faced by autonomous driving systems.

4.2.3 Wasserstein Distance:

The Wasserstein distance, often referred to as the "earth mover's distance," is a measure used to quantify the dissimilarity between two probability distributions. Conceptually, it represents the minimal amount of effort required to transform one probability distribution into another, envisioning distributions as piles of earth and the Wasserstein distance as the optimal transport

$$W(u,v) := \inf_{\pi \in \Gamma(u,v)} \int_{\mathbb{R}\times\mathbb{R}} |x-y| d\pi(x,y)$$

cost to reshape one pile into the other. Mathematically, the Wasserstein distance between two distributions, u and v, is formally defined as:

where Γ(u,v) is the set of all possible joint distributions (couplings) whose marginal distributions match u and v. This formulation enables the Wasserstein distance to capture both the shape and scale differences between distributions effectively, offering a more comprehensive measure of discrepancy compared to simpler metrics like mean absolute or squared differences.

In the context of fairness evaluation for pedestrian trajectory prediction models, the Wasserstein distance serves as a powerful tool for assessing prediction accuracy across various demographic groups. Specifically, it quantifies discrepancies in the error distributions of groups such as children versus adults, elderly versus adults, and men versus women, providing a nuanced understanding of how model predictions might systematically differ across these groups. Zhao (2021) established that the Wasserstein distance directly correlates with individual fairness, highlighting its ability to measure and reveal subtle fairness variations. By leveraging this metric, researchers can detect and quantify implicit biases embedded in prediction algorithms, guiding targeted efforts to enhance fairness in automated prediction systems deployed in safety-critical contexts like autonomous driving.

4.2.4 Algorithmic Out-of-distribution Evaluation Pipeline

Our evaluation pipeline consists of structured data partitioning and systematic model assessments. Each dataset undergoes track generation with a 50% overlap ratio for the train, validation, and test splits to enhance representational robustness. Models are trained uniformly across demographics to reflect realistic conditions. During testing, demographic-specific performance is analyzed independently.

For deterministic models, testing on a group that contains (n) test tracks results in the vector $(d\ =\ [d_1, d_2, \ldots, d_n])$, where d_n is the bounding box MSE on track n. For multimodal models, testing on a group that contains n test tracks results in the matrix $(E\ =\ [e_1, e_2, \ldots, e_n])$, where $(e_n\ =\ [e_{\{n1\}}, e_{\{n2\}}, \ldots, e_{\{n20\}}])$ is a vector containing the bounding box MSEs of 20 randomly sampled trajectory predictions on track n. Following standards set by previous studies [12, 28, 29], we use the best-of-20 approach for multimodal models by extracting the best prediction on each track, turning the matrix E in the vector $(m\ =\ [m_1, m_2, \ldots, m_n])$, where m_n is the minimum MSE value from e_n. Compared to deterministic models, the MSE values for multimodal models tend to be much lower because we essentially cherry-pick the most accurate trajectory prediction out of 20 predictions generated. Evaluations on JAAD and PIE use publicly available pretrained model checkpoints for consistency. For TITAN, models are trained anew with default hyperparameters due to the absence of established benchmarks.

4.3 Results and Discussion

Biased data often result in biased algorithmic outcomes, a concern increasingly recognized in machine learning research [22]. Finding unbiased and high-quality data continues to pose significant challenges, particularly in the context of pedestrian trajectory prediction, where representative demographic distributions are crucial. In our study, the datasets analyzed—JAAD, PIE, and TITAN—show a noticeable bias towards adult pedestrians. While some bias is expected, given the typical societal demographics in which adults outnumber children and the elderly, the representation of adults in our datasets is disproportionately higher than demographic statistics would suggest based on their respective filming locations. Table 1 provides a clear comparative overview of pedestrian age distributions within each dataset alongside the expected demographic distributions derived from official statistics.

Several plausible factors contribute to the observed demographic imbalance. Firstly, the filming times and locations inherently favor adult pedestrian representation; children are often in school during typical data collection hours, and elderly pedestrians tend to participate less frequently in walking activities, resulting in their underrepresentation in the recorded videos [23]. Additionally, annotator bias is likely another contributing factor. Pedestrians appearing further from the on-board camera present visual ambiguities, complicating accurate demographic annotation. Previous studies, including work by Wilson et al. [37], have demonstrated substantial inconsistencies among annotators in labeling subjective attributes such as skin tone. Age, similarly subjective and perhaps even more challenging to accurately determine visually, might lead annotators to default to labeling ambiguous pedestrians as adults, further inflating the adult representation.

4.3.1 Algorithmic Bias

Table 2: Model performance on different demographics in terms of mean MSE/CMSE/CFMSE.

		JAAD			PIE			TITAN		
Method	Group	MSE 0.5s / 1.0s / 1.5s	C_{MSE} 1.5s	CF_{MSE} 1.5s	MSE 0.5s / 1.0s / 1.5s	C_{MSE} 1.5s	CF_{MSE} 1.5s	MSE 0.5s / 1.0s / 1.5s	C_{MSE} 1.5s	CF_{MSE} 1.5s
BiTraP-D [40]	Child	**185** / **848** / **2826**	**2722**	**10945**	**85** / **468** / **1596**	**1572**	**7291**	289 / 1216 / **4218**	**4106**	**17161**
	Adult	182 / 662 / 2025	1900	7566	38 / 152 / 490	462	1880	360 / 1134 / 3120	2931	10421
	Elderly	147 / 410 / 1134	1040	4005	67 / 234 / 673	617	2346	**387** / **1374** / 4004	3773	13667
	Male	**191** / 611 / 1775	1641	6552	**43** / 160 / 490	454	1844	- / - / -	-	-
	Female	171 / **661** / **2064**	**1953**	**7806**	39 / **162** / **537**	**512**	**2127**	- / - / -	-	-
SGNet [34]	Child	147 / **736** / **2582**	**2481**	**10405**	**111** / **520** / **1581**	**1553**	**5940**	253 / 1033 / 3507	3418	13998
	Adult	**155** / 583 / 1844	1725	7177	33 / 133 / 449	422	1821	369 / 1174 / 3117	2945	10083
	Elderly	139 / 374 / 995	905	3479	61 / 202 / 586	533	2144	**392** / **1411** / **4243**	**4018**	**14698**
	Male	**166** / 545 / 1620	1494	6120	**38** / **141** / 449	416	1757	- / - / -	-	-
	Female	146 / **583** / **1876**	**1768**	**7236**	33 / 138 / **478**	**455**	**1976**	- / - /-	-	-
BiTraP-NP [40]	Child	**73** / 157 / 360	273	881	**43** / **173** / **457**	**415**	**1184**	158 / 296 / 566	509	1293
	Adult	72 / **167** / **390**	**303**	**993**	17 / 40 / 93	70	223	**187** / **356** / 697	554	1292
	Elderly	63 / 110 / 208	142	360	33 / 73 / 166	120	407	162 / 351 / **760**	**586**	**1572**
	Male	**78** / **161** / 341	247	740	**19** / **46** / **105**	76	253	- / - / -	-	-
	Female	64 / 157 / **385**	**307**	**1030**	16 / 40 / 97	**78**	**255**	- / - / -	-	-

Algorithmic bias refers specifically to biases introduced by the modeling algorithms themselves, independent of dataset biases. In our investigation, distinguishing algorithmic bias from dataset-induced bias remains complex since the models analyzed are trained on inherently biased datasets, as demonstrated previously. To explore algorithmic bias comprehensively, we assessed performance differences across demographic groups—age and gender—using three quantitative measures: mean squared error (MSE) presented in Table 2. Recognizing the complexity of isolating true algorithmic bias, future research initiatives aim to construct and utilize unbiased datasets explicitly designed to facilitate clearer assessments of model-induced discrepancies.

4.3.2 Age Bias

Our findings highlight pronounced discrepancies in the predictive performance of state-of-the-art trajectory models among different age groups. Generally, these models demonstrate significantly reduced accuracy when predicting trajectories for children and elderly pedestrians compared to adults, indicating potential vulnerabilities in current model designs and training paradigms.

1. **Results for BiTraP-D:** We found the mean MSE consistently highest among child pedestrians across all three datasets—JAAD, PIE, and TITAN. Compared to adult pedestrians, the average mean MSE for children was higher by approximately 23% on JAAD, 217% on PIE, and 25% on TITAN. Additionally, statistically significant differences emerged from Mann-Whitney U Tests comparing elderly and adult pedestrians, with all observed p-values well below 10^{-15} for both PIE and TITAN datasets, underscoring the robustness of observed performance discrepancies across age groups.
2. **Results for SGNet:** Our detailed evaluation of the SGNet model reveals noticeable discrepancies in predictive performance across different age groups. Specifically, SGNet exhibits the highest mean MSE for child pedestrians when tested on the JAAD and PIE datasets, and similarly, it yields the highest errors for elderly pedestrians on the TITAN dataset. On the PIE dataset, the differences in predictive accuracy are particularly striking: the mean MSE for child pedestrians is approximately 254% higher, and for elderly pedestrians, it is about 42% higher compared to adult pedestrians. The consistency of these performance discrepancies is further reinforced by the Mann-Whitney U Test outcomes, where the comparisons between elderly and adult pedestrian groups yield statistically significant p-values across both PIE and TITAN datasets. This robust statistical evidence underscores a pronounced bias of SGNet in accurately predicting trajectories for elderly individuals, potentially raising critical concerns for real-world implementations of autonomous driving systems.
3. **Results for BiTraP-NP:** Evaluating the BiTraP-NP multimodal model presents unique insights and some unexpected results. The highest mean MSE was recorded for child pedestrians on the PIE dataset, while elderly pedestrians showed the highest mean MSE on TITAN. Similar to the other models tested, BiTraP-NP exhibits statistically significant performance differences between elderly and adult pedestrians, clearly indicated by the significant Mann-Whitney U test p-values across the PIE and TITAN datasets. Interestingly, an anomalous finding arises from JAAD, where the highest mean MSE was surprisingly observed for adult pedestrians, contrary to results from other datasets and models. This unusual outcome is noteworthy, especially given the multimodal nature of the BiTraP-NP model, which selects the best prediction from 20 possible trajectories generated for each pedestrian track. Such an approach inherently biases the results toward lower errors, as it artificially selects the most accurate trajectory, a luxury unavailable in practical real-world autonomous driving scenarios. We recognize this as a methodological limitation and are currently investigating alternative strategies beyond the standard best-of-20 evaluation approach.

Despite notable differences in mean MSE values, statistical tests comparing children and adults frequently yield non-significant p-values, particularly on PIE and TITAN datasets. This is likely attributable to the limited representation of children within these datasets—children comprise only 0.9% and 1.4% of pedestrians in PIE and TITAN, respectively, as clearly demonstrated in Table 1. This small sample size diminishes the statistical power of our analyses, potentially obscuring meaningful differences. Nonetheless, a consistent trend emerges where the statistical significance of performance differences intensifies with longer prediction horizons. This indicates that as trajectory predictions become inherently more challenging, discrepancies in predictive accuracy between age groups become more pronounced.

An intriguing and initially counterintuitive result involves the strong predictive performance of all evaluated models on elderly pedestrians within the JAAD dataset. Initially, we hypothesized that this superior performance might result from annotation bias, as demographic labels (age and gender) in JAAD are available primarily for pedestrians close to the camera (approximately 25% of the dataset). It was assumed that training on unlabeled pedestrians located further away might artificially skew model performance. To test this hypothesis, we retrained all models using exclusively the subset of labeled pedestrians. Surprisingly, even under these controlled conditions, elderly pedestrians still exhibited superior predictive outcomes compared to both adults and children. Given this finding was not mirrored in the training data itself, we attribute this anomaly predominantly to statistical noise and specific sampling peculiarities within the JAAD dataset. Thus, while intriguing, these results should be cautiously interpreted, highlighting the necessity of careful dataset design and annotation procedures for reliable trajectory prediction modeling.

Table 3: Two - sided Mann-Whitney U Test p-values. Statistically significant p-values ($p < 0.05$) are in bold.

		JAAD	PIE	TITAN
Method	Demographics Compared	MSE 0.5s / 1.0s / 1.5s	MSE 0.5s / 1.0s / 1.5s	MSE 0.5s / 1.0s / 1.5s
BiTraP-D	Child|Adult	0.72 / 0.05 / **0.005**	0.84 / 0.83 / 0.48	0.16 / 0.12 / **0.008**
	Elderly|Adult	0.72 / **1e-4** / **1e-6**	**1e-77** / **5e-68** / **6e-55**	**2e-24** / **2-e18** / **9e-16**
	Male|Female	**0.02** / 0.11 / 0.43	0.96 / 0.81 / 0.42	- / - / -
SGNet	Child|Adult	0.54 / **0.03** / **0.003**	0.94 / 0.45 / 0.20	**8e-4** / 0.57 / 0.19
	Elderly|Adult	0.41 / **0.003** / **3e-6**	**2e-75** / **3e-65** / **6e-65**	**9e-27** / **2e-22** / **2e-20**
	Male|Female	**0.01** / 0.07 / 0.11	0.29 / 0.21 / 0.47	- / - / -
BiTraP-NP	Child|Adult	0.49 / **0.02** / **0.002**	0.92 / 0.48 / 0.41	**4e-6** / 0.06 / 0.31
	Adult|Elderly	0.10 / **0.02** / **2e-6**	**8e-93** / **8e-88** / **5e-81**	**9e-23** / **1e-23** / **9e-24**
	Male|Female	**0.01** / **0.01** / 0.05	0.56 / 0.54 / 0.21	- / - / -

To effectively mitigate the algorithmic bias, we further design a reliable fairness-aware mutli-modal learning paradigm to extract and fuse rich but noisy information from different objects (vehicles, infrastructures, and pedestrians), and jointly optimize the pedestrian detection and trajectory prediction accuracy across all pedestrian groups. A physics-informed multi-view graph will be designed to embed features extracted from different objects and modalities as ontological nodes and links the nodes with their physical interdependencies. We further explore Bayesian network-based methods and deep graph neural network approaches for pedestrian detection and prediction. Variational Bayesian inference that approximate the optimal sequential posteriors of pedestrian trajectory by jointly optimizing the marginal likelihood of multi-modal data, deep graph neural networks learn optimal node and edge embeddings for pedestrian detection and trajectory prediction. We further design adversarial on-manifold data augmentation scheme with approximated data manifolds of underrepresented pedestrians, and design fairness-aware loss function for aforementioned algorithms that jointly optimizes general accuracy and differences between group-level prediction accuracy across different modalities.

Table 4: Wasserstein distances for the MSE error distributions between different demographics. It's important to note once again that on JAAD, the performance of all three models is better on elderly than on adults.

Method	Demographics compared	JAAD	PIE	TITAN
		MSE	MSE	MSE
		0.5s / 1.0s / 1.5s	0.5s / 1.0s / 1.5s	0.5s / 1.0s / 1.5s
BiTraP-D [40]	child|adult	26 / 194 / 825	54 / 332 / 1136	159 / 449 / 1570
	elderly|adult	38 / 254 / 893	30 / 86 / 210	141 / 332 / 933
SGNet [34]	child|adult	14 / 159 / 763	81 / 396 / 1153	165 / 423 / 1182
	elderly|adult	21 / 209 / 849	28 / 69 / 148	158 / 414 / 1185
BiTraP-NP [40]	child|adult	7 / 54 / 205	29 / 118 / 332	114 / 158 / 289
	elderly|adult	12 / 55 / 176	16 / 35 / 81	79 / 133 / 166

5 CONCLUSIONS

This research presents a comprehensive and innovative fairness-aware, multi-modal pedestrian safety framework designed to leverage connected vehicle-infrastructure-pedestrian (VIP) networks, significantly enhancing safety and overall quality of experience for pedestrians. The primary objective of this work is to systematically identify, characterize, and mitigate algorithmic biases inherent in artificial intelligence (AI) models deployed within VIP systems. By utilizing advanced machine learning and representation learning methodologies, as well as analyzing both real-world video data and simulated scenarios, this study aims to ensure safety outcomes across diverse demographic groups.

The study is structured around three core research thrusts. Firstly, it focuses on uncovering and characterizing variations in pedestrian intent and addressing out-of-distribution (OOD) challenges. Through rigorous analysis of annotated datasets—JAAD, PIE, and TITAN—the research quantifies discrepancies in movement trajectories, walking speeds, and gait patterns across demographic groups defined by age, gender, and disability. Employing robust metrics such as hypothesis testing-based scores and Wasserstein distances, this phase effectively assesses and quantifies demographic discrepancies. The second thrust addresses OOD issues arising from limited data and distribution shifts that disproportionately impact vulnerable populations. To mitigate these challenges, fairness-aware, multi-modal learning algorithms employing distributional divergence metrics (e.g., Wasserstein distance, KL divergence) are developed. Advanced modeling approaches—including Bayesian inference, deep graph neural networks, adversarial on-manifold data augmentation, and fairness-focused loss functions—are utilized to improve pedestrian detection and trajectory prediction accuracy while minimizing discrepancies in model performance across groups.

Our comprehensive evaluation has clearly demonstrated that significant biases exist in state-of-the-art pedestrian trajectory prediction models, particularly with respect to age demographics. These models consistently produce higher prediction errors for children and elderly pedestrians compared to adults, highlighting substantial fairness and safety concerns for the deployment of these models in autonomous vehicle systems. While gender-based differences were minimal, the significant discrepancies observed across age groups call for targeted improvements in both data collection practices and algorithm design. Future work should focus on refining fairness metrics, developing unbiased datasets, exploring alternative multimodal prediction evaluation strategies, and actively mitigating identified biases to ensure reliable predictive performance for all pedestrian groups.

REFERENCES

1. National Highway Traffic Safety Administration. (2019). *Traffic Safety Facts 2017 Data – Pedestrians* (Technical Report, 216 pages). Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812681
2. Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management, 60*, 102387. https://doi.org/10.1016/j.ijinfomgt.2021.102387
3. Brandao, M. (2019). *Age and gender bias in pedestrian detection algorithms*. arXiv:1906.10490 [cs]. https://doi.org/10.48550/arXiv.1906.10490
4. Burstedde, C., Klauck, K., Schadschneider, A., & Zittartz, J. (2001). Simulation of pedestrian dynamics using a two-dimensional cellular automaton. *Physica A: Statistical Mechanics and its Applications, 295*(3), 507–525. https://doi.org/10.1016/S0378-4371(01)00141-8
5. Caesar, H., Bankiti, V., Lang, A. H., Vora, S., Liong, V. E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., & Beijbom, O. (2020). nuScenes: A Multimodal Dataset for Autonomous Driving. Retrieved from https://openaccess.thecvf.com/content_CVPR_2020/html/Caesar_nuScenes_A_Multimodal_Dataset_for_Autonomous_Driving_CVPR_2020_paper.html (pp. 11621–11631).
6. National Statistics Center. (2021). *Population Census 2020 Basic Complete Tabulation on Population and Households*. Retrieved from https://www.e-stat.go.jp/en/dbview?sid=0003445136
7. Chen, X., Xu, S., Han, J., Fu, H., Pi, X., Joe-Wong, C., Li, Y., Zhang, L., Noh, H. Y., & Zhang, P. (2020). Pas: Prediction-based actuation system for city-scale ridesharing vehicular mobile crowdsensing. *IEEE Internet of Things Journal, 7*(5), 3719–3734.
8. CIA. (2022). *Ukraine*. Retrieved from https://www.cia.gov/the-world-factbook/countries/ukraine/
9. Czech, P. (2017). Physically disabled pedestrians—Road users in terms of road accidents. In *Contemporary Challenges of Transport Systems and Traffic Engineering* (pp. 157–165). Springer.
10. Gitelman, V., Levi, S., Carmel, R., Korchatov, A., & Hakkert, S. (2019). Exploring patterns of child pedestrian behaviors at urban intersections. *Accident Analysis & Prevention, 122*, 36–47. https://doi.org/10.1016/j.aap.2018.09.031
11. Statistics Canada. (2017). *Census Profile, 2016 Census - Toronto, City [Census subdivision], Ontario and Toronto, Census division [Census division], Ontario*. Retrieved from https://www12.statcan.gc.ca/censusrecensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CSD&Code1=3520005&Geo2=CD&Code2=3520&SearchText=toronto
12. Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., & Alahi, A. (2018). Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks. Retrieved from https://openaccess.thecvf.com/content_cvpr_2018/html/Gupta_Social_GAN_Socially_CVPR_2018_paper.html (pp. 2255–2264).
13. Helbing, D., & Molnár, P. (1995). Social force model for pedestrian dynamics. *Physical Review E, 51*(5), 4282–4286. https://doi.org/10.1103/PhysRevE.51.4282
14. Herrero-Fernández, D., Macía-Guerrero, P., Silvano-Chaparro, L., Merino, L., & Jenchura, E. C. (2016). Risky behavior in young adult pedestrians: Personality determinants, correlates

with risk perception, and gender differences. *Transportation Research Part F: Traffic Psychology and Behaviour, 36*, 14–24. https://doi.org/10.1016/j.trf.2015.11.007

15. Hickox, K. L., Williams, N., Beck, L. F., Coleman, T., Fudenberg, J., Robinson, B., & Middaugh, J. (2014). Pedestrian Traffic Deaths Among Residents, Visitors, and Homeless Persons—Clark County, Nevada, 2008–2011. *MMWR Morbidity and Mortality Weekly Report, 63*(28), 597–602. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5779416/
16. Jeyachandran, S. (2020). Introducing the 5th-generation Waymo Driver: Informed by experience, designed for scale, engineered to tackle more environments. Waymo LLC.
17. Kogure, S., Watabe, K., Yamada, R., Aoki, Y., Nakamura, A., & Kataoka, H. (2022). Age Should Not Matter: Towards More Accurate Pedestrian Detection via Self-Training. Computer Sciences & Mathematics Forum, 3(1), 11. https://doi.org/10.3390/cmsf2022003011
18. Kotseruba, I., Rasouli, A., & Tsotsos, J. K. (2020). Joint Attention in Autonomous Driving (JAAD). Technical Report arXiv:1609.04741. https://doi.org/10.48550/arXiv.1609.04741
19. Lerner, A., Chrysanthou, Y., & Lischinski, D. (2007). Crowds by Example. Computer Graphics Forum, 26(3), 655–664. https://doi.org/10.1111/j.1467-8659.2007.01089.x
20. Maciag, M. (2014). Pedestrians Dying at Disproportionate Rates in America's Poorer Neighborhoods. Retrieved from https://www.governing.com/archive/gov-pedestrian-deaths-analysis.html
21. Malla, S., Dariush, B., & Choi, C. (2020). TITAN: Future Forecast Using Action Priors. Retrieved from https://openaccess.thecvf.com/content_CVPR_2020/html/Malla_TITAN_Future_Forecast_Using_Action_Priors_CVPR_2020_paper.html (pp. 11186–11196).
22. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A Survey on Bias and Fairness in Machine Learning. ACM Computing Surveys, 54(6), 115:1–115:35. https://doi.org/10.1145/3457607
23. Naumann, R. B., & Beck, L. F. (2013). Motor Vehicle Traffic-Related Pedestrian Deaths—United States, 2001–2010. MMWR Morbidity and Mortality Weekly Report, 62(15), 277–282. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4604973/
24. Navarro, P. J., Fernandez, C., Borraz, R., & Alonso, D. (2017). A machine learning approach to pedestrian detection for autonomous vehicles using high-definition 3D range data. Sensors, 17(1), 18.
25. Pellegrini, S., Ess, A., Schindler, K., & van Gool, L. (2009). You'll never walk alone: Modeling social behavior for multi-target tracking. Proceedings of IEEE 12th International Conference on Computer Vision, 261–268. https://doi.org/10.1109/ICCV.2009.5459260
26. Rasouli, A., Kotseruba, I., Kunic, T., & Tsotsos, J. K. (2019). PIE: A Large-Scale Dataset and Models for Pedestrian Intention Estimation and Trajectory Prediction. Retrieved from https://openaccess.thecvf.com/content_ICCV_2019/html/Rasouli_PIE_A_Large-Scale_Dataset_and_Models_for_Pedestrian_Intention_Estimation_ICCV_2019_paper.html (pp. 6262–6271).
27. Romero, S. (2022). Pedestrian Deaths Spike in U.S. as Reckless Driving Surges. The New York Times. Retrieved from https://www.nytimes.com/2022/02/14/us/pedestrian-deaths-pandemic.html
28. Sadeghian, A., Kosaraju, V., Sadeghian, A., Hirose, N., Rezatofighi, H., & Savarese, S. (2019). SoPhie: An Attentive GAN for Predicting Paths Compliant to Social and Physical

Constraints. Retrieved from https://openaccess.thecvf.com/content_CVPR_2019/html/Sadeghian_SoPhie_An_Attentive_GAN_for_Predicting_Paths_Compliant_to_Social_CVPR_2019_paper.html (pp. 1349–1358).

29. Salzmann, T., Ivanovic, B., Chakravarty, P., & Pavone, M. (2020). Trajectron++: Dynamically-Feasible Trajectory Forecasting with Heterogeneous Data. In Computer Vision – ECCV 2020 (Lecture Notes in Computer Science), 683–700. Springer. https://doi.org/10.1007/978-3-030-58523-5_40