Published: 26 April 2025

CHI 2025: CHI Conference on Human Factors in Computing Systems
April 26 - May 1, 2025
Yokohama, Japan

Conference Sponsors:
SIGCHI

RESEARCH-ARTICLE

ProxiCycle: Passively Mapping Cyclist Safety Using Smart Handlebars for Near-Miss Detection

JOSEPH BREDA, UW College of Engineering, Seattle, WA, United States

KEYU CHEN

THOMAS PLÖTZ, Georgia Institute of Technology, Atlanta, GA, United States

SHWETAK N. PATEL, UW College of Engineering, Seattle, WA, United States

Open Access Support provided by:

UW College of Engineering

Georgia Institute of Technology

CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (April 2025)
https://doi.org/10.1145/3706598.3713325
ISBN: 9798400713941

ProxiCycle: Passively Mapping Cyclist Safety Using Smart Handlebars for Near-Miss Detection

Joseph Breda
Paul G. Allen School of Computer Science & Engineering
University of Washington
Seattle, Washington, USA
joebreda@cs.washington.edu

Keyu Chen
Gridware
San Fancisco, California, USA
keyu.chen@gridware.io

Thomas Plötz
School of Interactive Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
thomas.ploetz@gatech.edu

Shwetak Patel
Paul G. Allen School of Computer Science & Engineering
University of Washington
Seattle, Washington, USA
shwetak@cs.washington.edu

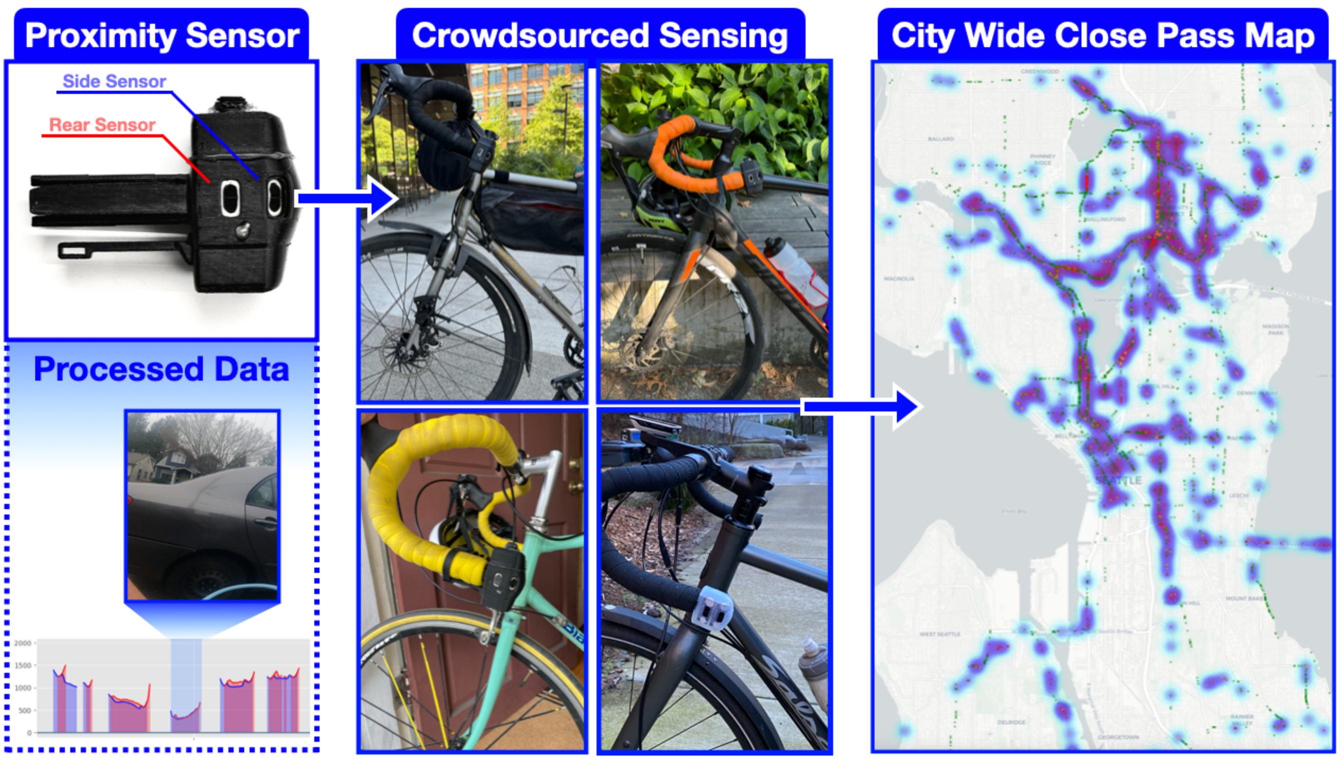

Figure 1: ProxiCycle leverages a dual proximity sensor design to retrofit and empower existing bike handlebars with the capacity to passively capture distance of passing cars during overtaking events. This data, when aggregated across many users, can provide useful insight into the safety of the road network without depending on users self-reporting.

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713325

Abstract

Active transportation is a valuable tool to prevent some of the most common causes of mortality worldwide, but is severely underutilized. The primary factors preventing cyclist adoption are safety concerns, specifically, the fear of collision from automobiles. One solution to address this concern is to direct cyclists to known safe routes to minimize risk and stress, thus making cycling more approachable. However, few localized safety priors are available, hindering safety based routing. Specifically, road user behavior is unknown. To address this issue, we develop a novel handlebar attachment to passively monitor the proximity of passing cars as a an indicator of cycling safety along historically traveled routes. We deploy this sensor with 15 experienced cyclists in a 2 month longitudinal study to source a citywide map of car passing distance. We then compare this signal to both historic collisions and perceived safety reported by experienced and inexperienced cyclists.

CCS Concepts

• **Computer systems organization** → **Embedded systems**; • **Human-centered computing** → **Mobile computing**; **Mobile devices**; **Smartphones**; *Ambient intelligence.*

Keywords

Smartphone, Bicycle, Urban Sensing, Mobile Sensing, Transportation, Safety, Public Health

ACM Reference Format:

Joseph Breda, Keyu Chen, Thomas Plötz, and Shwetak Patel. 2025. ProxiCycle: Passively Mapping Cyclist Safety Using Smart Handlebars for Near-Miss Detection. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3706598.3713325

1 Introduction

Cycling is a form of *active transport* which offers well known health and environmental benefits. In many industrialized nations the vast majority of adults do not meet physical activity guidelines to prevent chronic and fatal diseases such as cardiovascular disease or metabolic syndrome [73] which are responsible for the majority of mortality worldwide (approximately 35% of global fatalities [70]). Regular cycling and other forms of active transport have been shown to lower risk for cardiovascular disease, obesity, diabetes, some cancers, and even psychological disorders, demonstrating just some of the many health benefits [38, 45, 73].

Cycling as a mode of transport also has similarly positive impacts on environmental sustainability. For example, if the global population leveraged cycling for 2.6 kilometers (1.6 miles) a day (the current average in the Netherlands), global emissions from passenger vehicles would drop by 20% [8]. This is significant as transportation by personal passenger car is repeatedly found to be the largest single contributor to transportation sector's greenhouse gas (GHG) emissions [2] with the transportation sector being the largest portion of total GHG emissions worldwide [2, 32]. This motivates mode switching from car-based to cycling (or other active transit like walking or public transport) as one of the most direct approaches an individual can take to better their own health and reduce their environmental impact. The United Nations highlights this in goal 11 of "17 Sustainable Development Goals to combat climate change by 2030", which states improving the safety and broadening access to cycling (especially for vulnerable users such as women, children, and people with disabilities) as a primary contributor to sustainability [56, 60].

In order for populations to reap the economic, environmental, and health benefits of cycling, the majority of people need to feel safe doing it. Unfortunately, many people do not make this switch. For example, in the US alone, cycling trips make up less than 1% of all daily trips (for reference the leading country in cycling, the Netherlands, sees 25% of all trips including errands and commuting made by bicycle and 22% made on foot) [40]. Investigation into factors preventing cycling amongst Americans reveal that perceived safety, that is, how safe individuals feels irrespective of empirical measures of safety, is the main deterrent to the adoption of cycling, and specifically, the fear of injury or death due to collisions with cars [39, 60]. This belief is not unfounded as every year ≈ 41,000 cyclists die from road traffic crashes worldwide [60]. This number is disproportionately large when compared to total cycling trips. For example, in the UK in 2019, 6% of all road fatalities were people cycling despite cycling representing just 1% of all distance travelled and 2% of trips made.

Growing the cycling population is most directly achieved through modifying roadways to improve safety by separate cyclists from automobiles. For example, in Seville (Spain) the cycling population grew by 435% in one year – the same year the city expanded the protected bike lane network from 12 kilometers (7.5 miles) to 151 kilometers (94 miles) [49]. The addition of bicycle infrastructure indicates increased safety, which has been shown to be the deciding factor for many people adopting cycling [67]. However, changing the built environment takes time and political investment. So what about routes which do not yet have dedicated bike infrastructure? While more structural change is needed to make existing routes safer, one way cyclists can have immediately control over the safety of their cycling experience is to ride on known-to-be safer roads. Many experienced cyclists already do this is by developing their own preferred routes, defined by Kevin Lynch as *mental-maps* [47, 48] , by learning safety cues through trial and error. However, this approach is highly personal and developed through prolonged exposure to safety risks from riding on unsafe roads – something that many novice cyclists are not willing to do (i.e., explore-exploit dilemma). Instead of asking every cyclist to take on this up-front risk, a more accessible approach could be to aggregate existing knowledge of perceived safety captured by experienced cyclists and systematically communicate it to novice cyclists to help navigate them on safer and more comfortable routes as soon as their first ride. This can both mitigate safety risk by directing cyclists on safer routes and can lower the barrier of entry to many new cyclists, which can have further compounding effects on safety due to the principle of Safety in Numbers [31]. This measure of perceived safety, also referred to as level of stress (LOS) or level of traffic stress (LTS), [68] has been shown to influence people's decision to cycle over other modes, as well as increase comfort while riding [42, 67]. The perception of safety has a particularly strong impact on cycling adoption amongst parents and children [12, 16]. Prior work has called out a need for better quantitative measures of cycling perceived safety directly as a primary approach to promote cycling [18, 25, 26, 66].

However, existing systems for mapping safety and perceived safety are limited. For example, safety is measured primarily through historic bike collisions which are sparse and slow to reflect infrastructure changes [7]. Safety navigation therefore often defaults to simply reflecting the existence of bicycle infrastructure or lane width [20], lacking any further insight into other factors influencing safety or cyclist stress [48]. Observational studies or cyclist surveys aim to improve data quality [18], but like most survey methods, suffers from low coverage and high degree of aggregation [18, 26]. Platforms like Strava [80] have improved representativeness and granularity of cycling traffic data by involving citizens in open cycling data collection, but this paradigm has not yet been applied to safety [61]. Traditional car-based road safety literature addresses this using *collision surrogates* which are measurable events which are not, but correlate to, collisions such as *near misses* [44], but currently no scalable analog exists for bikes. Such a measure could serve as an indicator of collisions and therefore an indicator of empirical cycling safety. The current state of the art for sourcing bicycle near miss data is manually labeled video footage from constantly recording cameras such as the 100 Cyclist Project [46] or self-reporting through ad hoc web-portals like UpRide [1]. While these systems promise to leverage the scale of crowdsourced data to identify hotspots, they rely on manual reporting which does not scale. This is on top of the fact that cycling collisions themselves are already under-reported [22]. Not only is this tedious for users, but it is also prone to user bias as cameras do not capture true or standardized distance of near misses. Despite their flaws, the existence of these platforms and their user-bases motivates a more scalable bike safety sensing system.

To this end, we investigate a commonly reported unsafe (and uncomfortable) event in which a car overtakes a cyclist too close – a so-called "close pass". A close pass is when drivers overtake a cyclist on the road dangerously close, nearly colliding with or side-swiping the cyclist. In the field of injury prevention, Heinrich's Triangle proposes that these near-collisions are related to true collisions and avoiding them will in turn lead to a reduction in more severe collisions [28]. Additionally, these events influence cyclists perceived safety and in turn their willingness to ride [65, 68]. Many countries including the US, Germany, Belgium, and soon Japan [13, 14, 58, 78, 85] have legal distance margins that cars must provide when overtaking a cyclist, further grounding our definition of a close pass as a signal for safety. Over half of all US states have laws prohibiting drivers from passing cyclists within 1 meter (3 feet) as a safety measure, with as much as 1.5-2 meters common in some European countries [14, 58]. Despite this well-defined guideline, there is currently no scalable method for sensing and reporting where, when, and how severely drivers break these guidelines. To address this gap, we investigate the use of a bike mounted proximity sensor to automatically identify these close pass events passively while cyclists ride to provide a standardized, quantitative, and input-free method for sourcing a measure of cycling safety across the road network. The contributions of this paper are as follows:

(1) A formative study via a survey of cyclists of varying expertise to understand the potential efficacy of empirical safety data in facilitating cycling.
(2) The design of a smart proximity sensing handlebar plug and technical validation demonstrating the feasibility of passively measuring the distance of passing cars as cyclists ride, and
(3) A 2-month longitudinal deployment with 15 local cyclists to develop an aggregate map of observed close passes as a measure of cyclist safety which we then validate against historic collision data and perceived safety surveys.

ProxiCycle advances the line of inquiry on bicycle safety across the road network, motivated by the current lack existing approaches for measuring bicycle safety in a standardized manner at scale. We draw on traditional automotive collision surrogate literature [44] and existing bicycle crowdsourcing systems [10, 23, 80] when designing ProxiCycle to be scalable, low-cost, fully passive, and easy to deploy. Our system contributes to the growing field of smart active transportation and bicycle HCI [3, 15, 17, 51, 52, 63, 66, 69, 75].

2 Related Work

In this section, we will discuss the prior work in traditional safety sensing through collision surrogates and technology for urban mobility.

2.1 Traditional Collision Surrogates

Traditional road safety literature addresses the scarcity of historic collision data through collision surrogates which are measurable events which correlate highly with or even predict collisions. Some examples of existing collisions surrogates include *Near Misses*– when vehicles nearly but do not collide– which can be extracted from traffic camera footage [74] or self-reported on crowdsourcing platforms like UpRide [1] as well as *hard-braking events*–when a vehicle stops abruptly– which can be detected by accelerometers either built into or coupled to cars (i.e., smartphone mounted on the dashboard) [44]. Similar paradigms have been used to to explore bicycle swerving in response to driver behavior as a potential surrogate such as in CycleSense. However, the authors note that incident types such as tailgating or close passes may go undetected by motion based models as cyclists may not change their motion profile due to these dangerous events. Bicycle motion is notably prone to false positives due to motion artifacts such as braking and riding on uneven terrain as indicated by CycleSense's reported precision of 0.035 [35, 36]. Similarly, Apple Watch's accelerometer based crash detection systems [4] experienced high false positive rates due to the indirectness of the surrogate signal. For example, in January 2023, Apple's crash detection system triggered 134 false emergency calls in Hida Japan's mountain region (more than 14% of all emergency dispatch calls) reportedly due to iPhone 14s triggering while user's were skiing [57]. These false positives may be mitigated by deploying more accurate sensing systems which either: (1) include context signals to identify situations which can trigger false positives; or (2) leverage sensors which measure events which are more directly associated with the incident being detected.

2.2 Augmented Cyclist Perception

Prior work has also explored novel sensing approaches for sensing around the cyclist to extend their perception. For example, CycleGuard leverages a smartphone sonar attachment to detect and alert riders of right-hook interactions at intersections–when a driver cuts-off a cyclist while taking a right hand turn [33]. HindSight leverages a 360 degree camera mounted on a cyclist's helmet along with a bone conductance speaker to provide sonified feedback to extend spatial awareness beyond the capacity of human biology [76]. Similarly, Velo.ai launched an AI powered tail light camera for identifying approaching vehicles [83], but cannot measure proximity of passing cars directly. The Garmin Varia tail light radar [21] addresses this by leveraging approach distance vehicles from the rear, but does not measure the passing distance during overtaking events.

While these systems can identify approaching vehicles and can communicate this information to riders through alerts, it is unclear whether these alerts arrive within the critical time window, allowing riders to take evasive action. Specifically, vehicle operators have a perception-response time of 1.6 seconds before they can be reasonably expected to respond to visual feedback [59]. While this may not seem like a long time, the average cyclist rides at approximately 9-15 mph [37] with cars typically passing 10-15 mph faster. This means a feedback system would need to detect approaching vehicles over 10 meters away. At such long distances, it may not be possible to predict whether a car will pass too close or provide a safe amount of space, potentially rendering such an alert useless. They can also be potentially distracting or sources of stress which has been shown to further impede cyclist safety [50, 71]. Furthermore, camera and radar based systems like this can be expensive (Velo.ai at $400 and Garmin Varia at $200 USD respectively), making it unlikely for such a system to scale enough to impact the safety of the broader population such as novice cyclists which have the most to gain from safety technology.

2.3 Crowdsourcing and Passively Sensing City Scale Informatics

Prior work has addressed the need for scaling data to population levels with crowdsourcing or passive sensing to map the build environment similar to how cyclists build up mental-map of safety through experience and exploration [48]. However, by leveraging sensing or crowdsourcing, these maps can be aggregated across the population and shared with new users who would previously need to gather this information themselves. For example, Project Sidewalk combines crowdsourced labels and existing Google Street View (GSV) images to identify sidewalk accessibility issues to facilitate better navigation and awareness around sidewalk accessibility hazards [27, 43, 72]. crowdsourcing alongside GSV has been used for other mapping task such as mapping street level objects [54] or even perceived beauty of the built environment [64] to improve city scale monitoring and navigation. Alternative signals such as geolocated social media posts have been used to map urban well-being [62]. Prior work has also demonstrated the feasibility of deploying passive sensing for sourcing entirely new datasets such as city-wide air quality by mounting sensors on vehicles like garbage trucks or passenger cars [5, 6, 55]. Similarly, passively monitored GPS data from smartphones has been used to naturallistically identify disease spread hot spots or influences of the built environment on people's health [53, 81].

We draw on this body of city scale mapping work to design our system which address the unique gap in existing bicycle technology for sensing collision surrogates which is: inexpensive, easy to manufacture, distribute, and deploy, and is unobtrusive and fully passive. ProxiCycle is therefore positioned as a "plug-and-play" smart handlebar plug which is inserted and then forgotten about while continuously recording safety metrics across the population of cyclists, thus allowing safer routes to emerge in the data naturalistically.

3 Formative Study

We conducted a brief formative study through a locally distributed fully anonymous IRB approved survey to understand the potential affordances of a map of cycling safety. This survey was designed to answer the following questions: (1) what do cyclists consider when making the choice of whether and where to cycle, (2) how might a map of perceived safety impact these choices, and (3) do answers to these questions differ across cyclists of varying experience?

3.1 Survey Structure

We recruited survey participants through advertisement within the university, local neighborhood newsletters, social media messaging, and snowball sampling. After 3 weeks we received 389 responses. Participants were asked demographics questions and then to self-report their experience with cycling on a scale of 1 = very inexperienced to 7 = very experienced (participants were prompted that experience was defined as combination of frequency and consistency cycling). We then asked a series of five questions probing their perception of the safety of cycling, what factors influence their perception, and whether they would leverage knowledge of safety across the road network to influence their choice to cycle more. A full list of survey questions as they appeared in the survey is included in the Appendix A. To prioritize recruiting a large survey population, we designed the survey to be completed quickly utilizing a combination of multiple choice and Likert scale questions. Participants were also given an optional short answer response for each question to further elaborate on their answer.

3.2 Survey Responses

Of the 389 responses, 71% self-reported above average experience cycling (experience 5-7 out of 7) indicating our sample population tends towards experienced cyclists. We discuss responses to each question in turn below. In the following discussion we provide quotes from the open response questions along with the survey participant ID and their self-reported cycling experience rating.

Perceived Danger of Cycling: The mean response was 3.61 out of 5 with less experienced riders (experience 1-3 out of 7) ranking cycling as more dangerous than more experienced cyclists (experience 4-7 out of 7). Less experienced riders (experience 3 out of 7) ranking it the highest at 3.91. These results are illustrated in Figure 2. About a quarter of responses used the open response section to elaborate that cycling itself is not dangerous but rather having to share the road with cars is. One participant said *"the danger is outside of my control. My biking is safe, it's others that make it dangerous."* (S27 6/7). Other respondents shared similar sentiments that: *"The danger from cycling comes from cars."* (S5 6/7), and *"Cycling as a mode of transit is very safe. Traffic is not safe but that is not cycling's problem"* (S165 7/7). Some participants even brought up the varying safety across the road network as a result of road-specific infrastructure and driver behavior – a key motivation for this work. One participant said *"depends on route/infrastructure, and how accustomed drivers are to being around bikes"* (S35 7/7), while another mentioned *"I work very hard to avoid high traffic volumes and high speed [roads], and because of that I feel mostly safe. During most trips, however, I have to face at least one unsafe segment or intersection."* (S122 7/7). Some participants went further and noted

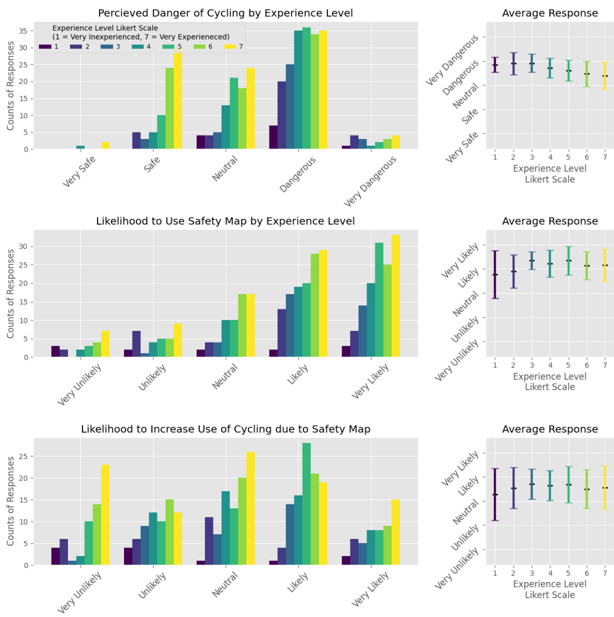

Figure 2: Three grouped histograms showing the count of responses for the three core survey questions, grouped by the self-reported experience level of respondent on a 7-point Likert scale (from 1 = very inexperienced to 7 = very experienced). The mean and standard deviation for each question by experience level are shown to the right. The histograms show, from top to bottom, the respondent's perceived danger of cycling, their likelihood to use a mapping application which communicates variable safety across the road network, and their likelihood to increase their utilization of cycling for transportation as a result of this mapping tool.

that knowledge of varying safety across the road network can help improve their experience cycling. One notes, *"Knowledge of streets and traffic patterns is essential, but once you have a familiar and well considered route, it is very safe"* (S245 7/7), and another says *"You do have to know which streets to ride on and ride defensively"* (S95 7/7). Another participant highlights the frustration of feeling forced to do this legwork, motivating a platform to share this information across the cycling population: *"The single worst part about cycling here is [those who ride] bike or walk either have to research their route or risk being thrown into active traffic."* (S223 3/7) Participants like this directly cite the barrier of entry to novice cyclists. These results indicate that less experienced cyclists perceive cycling as more dangerous and affirm our expectation that experienced cyclists may ride on a safer road network by leveraging prior knowledge to avoid dangerous streets (i.e., mental maps accumulated through experience) – something that we aim to share with the broader cycling population.

Ranked factors encouraging and discouraging cycling: Next, participants were asked to rank factors which encourage or discourage cycling. The ranking stratified by respondent's cycling experience is shown in Figure 3. Among the factors which encourage cycling, access to infrastructure and safety from traffic were

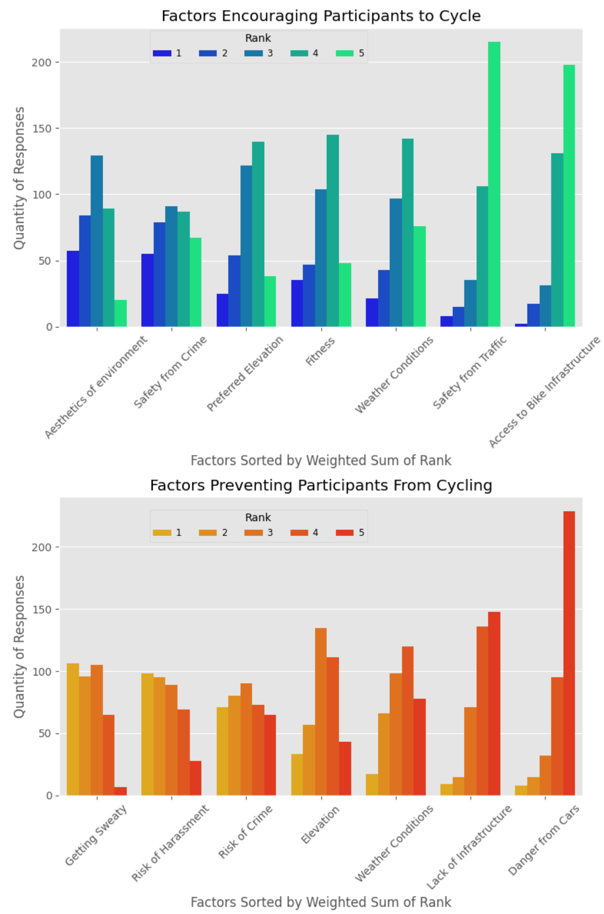

Figure 3: Two grouped histograms showing the ranking of influence (from 1 being lowest to 5 being highest) for each factor encouraging (top) and discouraging (bottom) respondents from cycling. Each histogram shows the total number of respondents who assigned each rank to each factor. The factors are sorted in order of weighted average rank from lowest (left) to highest (right) influence.

ranked most highly. Among factors which prevent cycling, danger from cars was ranked most highly by a significant margin across all experience levels. A few participants cited personal experiences with close-calls or collisions which have since detured them from cycling *"having had plenty of close calls with inattentive drivers I have little interest in riding on roads, or even unprotected bike lanes"* (S340 4/7). This result indicates that conflict with cars is a significant factor preventing cyclist adoption.

Likelihood to use safety map and perceived effect on cycling: Respondents indicated they would be very likely to use a safety based navigation map with an average response of 4.36 out of 5. Those with intermediate levels of cycling (3-5 out of 7) were most likely to use it. These results are illustrated in the top row of Figure 2. Finally, when asked how likely they were to increase their frequency of cycling given access to this data, Figure 2 shows respondents overall were more likely to increase their cycling with a mean response of 3.62 out of 5. The mean response was greatest for below average experience (3 out of 7) cyclists at 3.73 which intuitively makes sense as the most experienced cyclists may already cycle at max capacity or feel secure with their "mental map" of safety given their increased experience [48]. This intuition is reinforced by short-answer responses from expert cyclists (7 out of 7) participants who note, *"A map will not help me cycle more. I already use my own senses and memory and find the best routes, even for places I've never been"* (S44 7/7). On the other end, those with no experience at all may not have interest in cycling. Though some inexperienced cyclists may want to cycle more but feel blocked by the complexity of route planning: *"Confidence in navigating safe cycling routes is my #1 barrier to cycling more. I don't feel confident that Google Maps or other navigation systems would provide a safe route and despite my relatively high awareness of the Seattle bike lane system, there is still a prohibitive amount of route planning research for casual use"* (S365 3/7). Quotes like these demonstrate how experienced cyclists tend to identify these safer routes through trial and error, however, this may not be an option for inexperienced cyclists or even some experts who are not willing to make the safety trade off by cycling unfamiliar territory [48]. This is directly called out by a novice cyclists (1 out of 7) who says *"My iPhone and the Maps app fundamentally changed riding transit for me, an always up to date bike map with honest info would be amazing"* (S146 1/7). That being said, some expert cyclists noted this may still be useful in unfamiliar environments *"I know my way around my home well, but a map of that type would help in places I don't know as well or am new too"* (S139 6/7). These findings further motivate ProxiCycle by identifying that access to this quantified measure of safety is useful to cyclists of all levels, but specifically can make cycling more approachable by increasing the likelihood of cycling among less experienced riders.

4 System Design

In this section we will discuss our design considerations when building ProxiCycle as well as the physical and software implementation. We finish this section with a brief road map of the study design from technical evaluation, to deployment, and finally spatial analysis of the crowdsourced map of close passes.

4.1 Design Consideration for a Collision Surrogate Sensing System

Our goal with this work is to provide a scalable measure of safety across the road network which can be used to navigate cyclists on safer routes, in turn lowering the barrier of entry to cycling and growing the cycling population. Following traditional injury prevention and collision surrogates research [28], we designed ProxiCycle to measure safety passively as users ride through the lens of close passes as a potential collision surrogate. A strong collision surrogate will occur much more frequently than actual collisions and represent not just the most severe collisions, but all possible collisions, including those which may go unreported. This approach addresses the scarcity of safety data as close passes occur much more frequently than self-reported near misses and collisions and eliminates both user burden and user bias [24].

This decision to use close passes as a collision surrogate is supported by existing work using lane width as an indicator of LTS by urban planners as well as legal passing margins referenced in Section 1 [29, 68]. We made this decision despite right-hook collisions at intersections making up the majority of reported bicycle collisions [30] as studies show that severe under-reporting has introduced strong bias into the types of collisions which are available for analysis in favor or the most severe (i.e., those resulting a visit to the emergency room) [22]. Conversely, a naturalistic observation study of risk-factors of road cycling found that *side-swipes* (where a car collides with a cyclist side-to-side while passing) were the most commonly observed hazard, more than twice as common as *right-hooks* (the most commonly reported collision, where a car performs a right turn into the cyclists path) and over 10x more common than *dooring* (where a parked car door is opened in the path of a cyclist) [34]. Finally, prior work indicates that ride quality is improved when riders are most present and less distracted by technology [71], so we constrain our design to: (1) work entirely automatically with no user input, and (2) for our device to be as physically unobtrusive as possible.

4.2 Proximity Sensing Technology

The most popular off-the-shelf proximity sensors include: acoustic (sonar), radar, stereoscopic video, and light based time-of-flight (ToF) sensors. Single lens cameras are sometimes used to estimate distance but these approaches are error prone. In this study we chose to leverage the ST VL53L8 multizone ToF sensor [79] for its small size, interpretable output signal, minimal data overhead (on the order of kilobytes per minute) and lack of multipath issues (common with both radar and acoustic ranging). The narrow field-of-view of ToF sensors allows for the system to easily isolate passing cars to the left of the cyclist without interference from other objects in traffic (i.e., parked cars to the right or vehicles traveling in front of or behind cyclists).

The VL53L8 works by sending an infrared light pulse (Tx) in a single direction and listening for a response (Rx) due to reflections off surfaces in range and in the field-of-view. Unlike many ToF sensors, the VL53L8 is specifically designed for outdoor use making it more robust to infrared interference from the sun. ToF sensors are also not commonly used for automobile proximity sensors so there is low risk for interference from other vehicles on the road. Anecdotally, we experimented with ultrasonic acoustic ranging sensors and found they suffered from interference from existing ultrasonic noise on the road. This was likely due to sensors built into nearby vehicles for blind spot detection. The small size of data produced by the VL53L8 also allows for easy communication over I2C and BLE, even at high data rates. These ToF sensors are also smaller than parallax based depth camera alternatives, allowing for multiple sensors to be integrated onto the same device without compromising the unobtrusive form factor and do not suffer from self-interference like sonar and radar might. Finally, the data from these sensors do not produce visual features and are inherently privacy preserving.

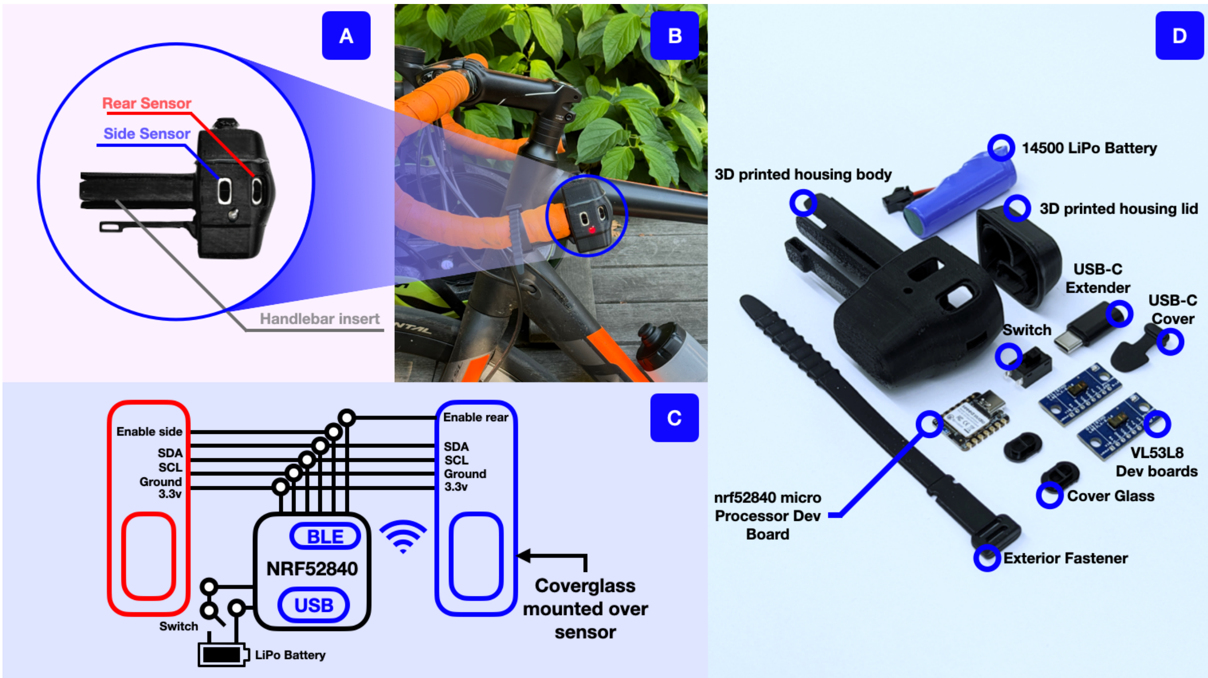

Figure 4: A) Close up of the fully fabricated Smart Handlebar Plug. B) The placement of the device on the end of a bicycle handlebar. C) A circuit diagram illustrating how components are connected. D) A layout of components to fabricate the device.

4.3 Hardware Design & Placement

Physical Placement: When designing our custom sensing platform we constrained our sensor's physical design to: (1) preserve limited handlebar real estate, and (2) position the sensor outside the rider's grip position to avoid measurement noise from the rider's hand or arm blocking the sensor. This led us to the final design of a smart handlebar plug by integrating the proximity sensor into the plug used to secure handlebar tape at the end of the handlebar. While the current prototype is designed for drop bar handlebars, similar designs can exist for flat bars or other handlebar shapes so long as the designs maintain an unobstructed field-of-view.

Our sensor is mounted by sliding it into the open end of the hollow handle bars. This design can be seen in sub-figure A) and B) of Figure 4. Given the small size of these sensors, we were able to integrate both a side facing sensor and a rear facing sensor to understand the velocity of passing objects based on the order of sensor triggers. This design avoids consuming handlebar real estate, leaving room for other devices such as bike computers, phones, or bells to be mounted on the handlebars[1] while also placing the sensor as far to the left of the bike as possible, which simultaneously increases the sensor range with respect to the bike and minimizes the likelihood of the rider's arms interfering with the sensor's field-of-view. The whole device including 3D printed housing is 6cm by 4cm by 3cm and was primarily constrained by the 14500 battery used for this prototype. However, we did not prioritize size optimization in this work and the majority of the housing is empty space to accommodate future development with additional components. The components themselves are 2.7cm by 2.7cm by 0.9cm and can be further reduced using a custom PCB.

Hardware Components: The sensor data is captured from 2 VL53L8 proximity sensors and communicated to a generic microprocessor (NRF52840) on the Seeed Studio Xiao development board over I2C. The Xiao comes complete with a Bluetooth Low-Energy (BLE) antenna used to transmit the data to a paired smartphone. The system is powered by a 3.7v 1000mAh 14500 LiPo cell battery to maximize battery life while continuously collecting raw data. The components of the device and block circuit diagram can be seen in sub-figures C) and D) of Figure 4. Each of the VL53L8 sensors in continuous mode have a current draw of 40mAh [79] and thus a 1,000mAh battery will last over 12 hours powering both sensors in ideal conditions. This is a reasonable battery life as cyclists often use bike lights which typically have < 8 hours battery life. The battery capacity, and therefore size, may be significantly reduced once power optimized for on-device processing and limiting wireless data transmission. The Xiao also natively supports power over USB-C allowing the device to be powered by external battery packs, dynamo motors, e-bike charging ports, or even the smartphone itself. The whole system can cost less than $25 to build in house using off-the-shelf components which is already an order of magnitude cheaper than the closest alternative (i.e., Garmin Varia $200). The fabrication cost (i.e., PCB, plastic housing, and assembly) could be further reduced with manufacturing at scale.

[1]This proved significant in our user-study as the majority of our participants had fully instrumented their handlebars leaving no room for additional attachments. This design may therefore improve the usability of the device.

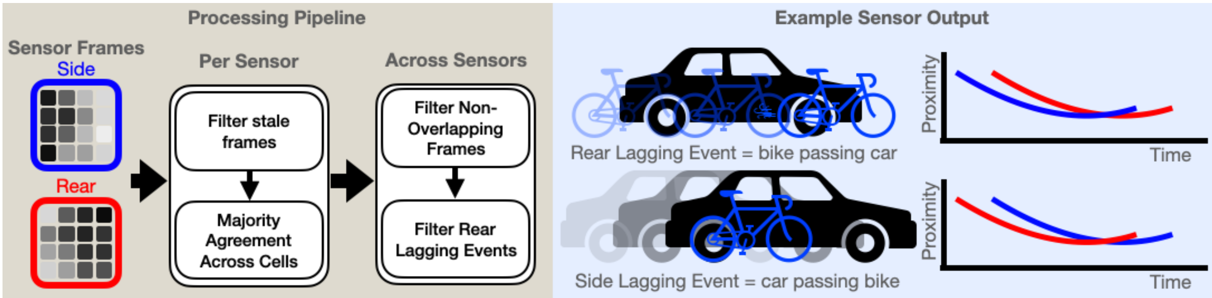

Figure 5: Flow diagram of our signal processing procedure for data recorded using our custom sensor platform for close pass event recognition. At each frame of each of the two sensors the system checks if the majority of cells updated agree within some threshold τ of each other (i.e., most cells update to a similar value at the same time step) to identify sensor events. Then the system checks if there is overlap between sensor events from each sensors and logs a close pass if the rear sensor event triggers prior to a side sensor event indicating that the passing object has a positive relative velocity with respect to the bike.

4.4 Signal Processing

To process the readings we employ a signal processing chain shown in Figure 5. The raw data from the VL53L8 sensors are low-resolution 4x4 frames representing the depth at each of the 16 cells within the sensor's field-of-view at the time a frame was captured. Each sensor samples at 10Hz including BLE transmission delay. To convert this raw timeseries data into meaningful sensor events, we first filter out any cells unchanged from the previous frame (which occurs when no objects are in range for that cell). We then log a sensor events if the majority of cells in the sensor frame agree to the same value within a threshold parameter, which we set as 0.2 meter empirically through early experimentation during development. This ensures that sensor events are only logged when large objects like cars block the sensor's field-of-view , i.e., one or more small objects like branches or distance pedestrians in the sensor's field-of-view will not trigger this condition. While there is no guarantee that this approach strictly identifies cars, we empirically saw during testing that passing cyclists and pedestrians rarely satisfied this condition at the typical passing distance due to the wide field-of-view of the VL53L8. The 0.2 meter threshold was used to give a small tolerance to natural noise between cell's measured distances or reflections off surfaces with modest curves while still maintaining agreement between cells. We then log events as a passing event if both the rear and side sensors trigger events overlapping in time and the rear sensor is triggered first. This lag in the side sensor indicates that an object passing the bike approached from the rear with a positive relative velocity with respect to the bike. Events where the rear sensor lags behind the side sensor indicate a negative relative velocity, i.e., cyclist passing a static object or slower road user. Since cars can pass cyclists with variable velocity, we did not threshold the magnitude of the lag between sensors. This ensures that cars would be detected, irrespective of their speed. This relationship is illustrated in Figure 5. We then compress the 16 cell ToF sensor timeseries signature to a single pass distance using the inner-quantile mean (IQM) of the side sensor at each frame and select the minimum IQM across the duration of the pass. We then filter out any pass distance above that of a large road lane at 3 meters and below 0.2 meters as these short distance triggers are due to cyclist hands during grip adjustments. We use the legal definition of a close pass as 1 meter and consider sensor events at the 1-3 meter range to be safe passes. In practice, we identified borderline cases where 1.3 meters felt uncomfortable to our participants, and for the purposes of evaluation, we report close pass and borderline close passes (<1.3 meters) as positive. This is closer to the most progressive legal passing margin in the US at roughly 1.2 meters (4 feet) which was written into law the same year this paper was written [11].

4.5 Study Road Map

Figure 6 shows a road map of ProxiCycle evaluation. After completing hardware development, we conduct a technical evaluation in Section 5 to validate the signal processing discussed above in Section 4.4. We then conduct a longitudinal deployment which we discuss in Section 6 to crowdsource a city-scale dataset of close passes. We finally show that this map of close passes can be used to analyze safety across the road network in Section 6.1.

5 Technical Evaluation

In this section we conduct two technical evaluations. First, we validate the chosen sensor modality is capable of measuring proximity accurately in the specific handlebar position and context of both moving bike and moving cars in multiple weather and ambient lighting conditions. This validation is essential to ensure trustworthy readings during in-the-wild deployments where it is not feasible to collect ground truth distance. Next, we investigate potential sources of error in close pass detection through a 7-user IRB approved pilot study where we compare triggered sensor events against manually reviewed simultaneously captured video footage on unrestricted routes.

5.1 Experiment 1: In Lab Validation

As a proof-of-concept we investigate the accuracy of the VL53L8 when mounted on the left stem of the bike handlebars[2] for sensing the proximity of passing cars through a controlled in-lab experiment. To ensure the results of this experiment most directly reflect

[2]We note that all user study trials were conducted on roads with right lane driving where bike lanes were always right of the nearest traffic lane. We note that this set up fails to capture close passes to the right of the cyclist. This system can easily be extend to capture both sides by adding a second device.

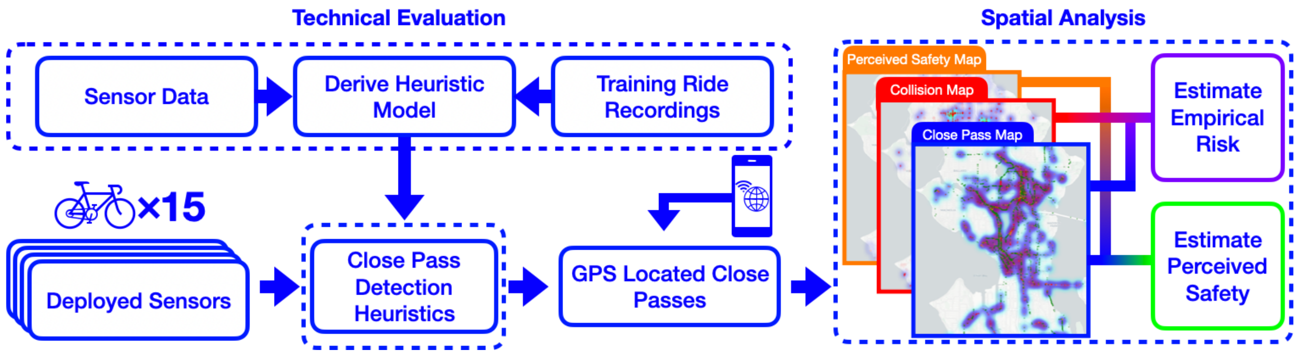

Figure 6: An overview of the study road map. The study begins with a technical elevation to derive the signal processing technique using raw sensor data compared to video recordings. The we leverage this model during a longitudinal sensor deployment to source GPS located close passes which are used in a spatial analysis of bicycling safety across the traveled routes.

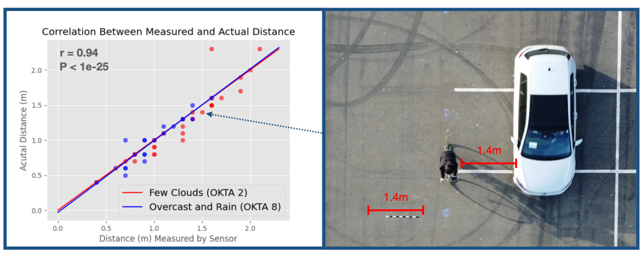

Figure 7: (Left) The correlation (Pearson's r of 0.94 and P-value < 1e-25) between ground truth distance measured manually in pixels from aerial view using reference marker and ProxiCycle's distance readings captured on both sunny and rainy days. (Right) A frame from a drone video illustrating how ground truth distance was captured using the reference marker in frame.

the real world interaction, we chose to simulate passes with both a moving bicycle and moving car.

We used photogrammetry (i.e., simultaneous video capture) to capture a reliably ground truth distance to compare to our sensor readings. To do this, we leveraged a 1 meter stick painted alternating black and white every 10 centimeters in frame as a visual reference of pixel count to distance in our video. We found it was easiest to consistently capture space between the bicycle and car from an aerial perspective (as opposed to a tripod on the ground in-line with the pass). So, we recorded video from a DJI mini 2 drone in hover mode directly above to simulated passes. By capturing video of both the simulated pass and black and white reference marker, we were able to manually map pixel quantity of each 10cm subsection of the marker to the real world distance at the 10cm resolution. Figure 7 shows the length of the reference marker compared to the length of a pass in the same frame. Because there was potential for the altitude of the drone to shift during flight, the mapping of pixel to distance was recalculated for each pass (though in practice it remained relatively constant). A sync gesture visible from both the camera and the sensor was performed at the start of each data collection.

By flying the drone at a sufficient altitude and centering the location of simulated close passes in frame, we can minimize any possible effects from edge distortion (e.g. barrel distortion) due to curvature of the lens. It is not essential to use a drone, as a similar approach could be replicated using a camera mounted from a tall structure such as a building given proper orientation and line-of-sight. For all trials both the cyclist and car were moving in a straight line to simulate the most common approach of automobiles overtaking a cyclist. the car was traveling between 24 and 48 kilometers per hour. To assess the effects of weather (both lighting conditions and precipitation) on sensor readings, we repeated the same experimental procedure twice: once on a mostly sunny day with few clouds and again on an overcast rainy day. The cloud cover on these days were recorded using OKTA scale, a meteorological standard metric, and were recorded as OKTA 2 and 8 respectively.

We captured 42 simulated rectilinear passes at variable distances from 0.4 meters to 2.2 meters and found our sensor readings to have high correlation with actual distances with an aggregate Pearson's r of 0.94 (0.93 in sun and 0.95 in rain). This relationship is statistically significant with an aggregate P-value of 1e-25 (1e-13 in sun and 1e-11 in rain) indicating that the proximity metric captured from our smart handlebar plug is highly correlated with the ground truth distance with no significant variation due to lighting conditions or precipitation. These results can be seen in Figure 7 alongside a example frame from our ground truth recordings.

5.2 Experiment 2: In the Wild Pilot Study

Next we deployed our system in-the-wild with a group of 7 pilot riders (each riding their own personal bicycle) in Seattle to evaluate the detection accuracy of our system at identifying cars against other real-world confounds. We deployed each sensor module in conjunction with a GoPro action camera mounted on the handlebars pointed perpendicular (left) to the bicycle to record passing objects during each user's ride. In the video we capture all instances of passing objects including cars and other large objects which trigger our sensing system. A sync gesture visible to both the sensor and camera was performed at the start of each ride. The synchronized video and sensor feed were manually reviewed and annotated for evaluation.

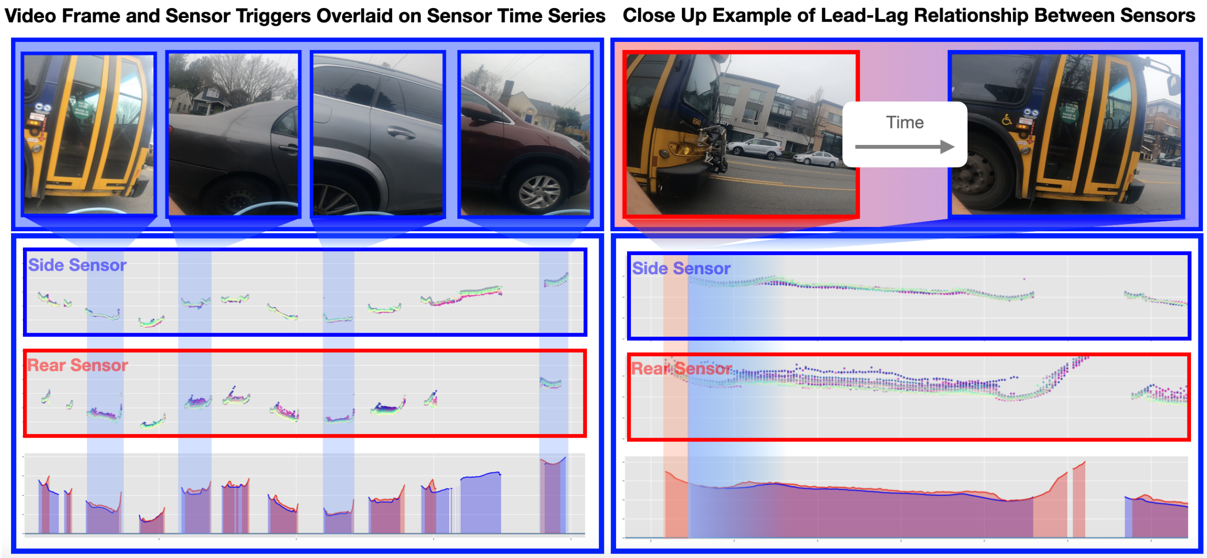

Figure 8: (Left) A visualization of the video capture overlaid with processed sensor triggers. (Right) A close up example of the lead-lag relationship used to distinguish passing vehicles from instances of the cyclist passing something in the environment.

Our participants were recruited through social media messaging and word of mouth from a local bike commuting community. All participants road their own bicycle retrofit with our sensor for the study. Self-reported experience level of our riders varied with: 4 cycling a few times a month, 2 cycling multiple times a weekly, and 1 who cycled multiple times per day. Each user road with our sensor and GoPro for 45 minutes to 1.5 hours (limited primarily by the GoPro battery life). The GoPro was mounted on the left side of the handlebars facing left (i.e. towards passing traffic). The routes were not pre-defined to preserve the naturalistic discovery of possible sources of error. Riders primarily chose familiar routes during their rides. The rides were carried out on different days with varying weather conditions including light rain, clear skies, and overcast skies. All trials were conducted during daylight hours to ensure the video feed from the GoPro was interpretable.

To generate ground truth labels for evaluating ProxiCycle's performance, the accompanying video footage was manually annotated.[3] Because of the lack of ambiguity in these labels, i.e., car pass or not, all videos were annotated by a single annotator over multiple days in hour long sessions with breaks in between sessions to avoid fatigue. Annotations were made by scrubbing through the videos one-by-one from start to finish and recording the timestamps where cars passed the cyclist. This was done over two passes. First, an initial blind pass, i.e., without access to ProxiCycle sensor data, was made to visually annotate every time a passenger automobile, i.e., car, truck, bus, etc., passed the cyclist in the footage. Then a second pass of each video was made by the same annotator, this time with access to the sensor events recorded by our signal processing technique defined in Section 4.4 to compare ProxiCycle against the annotations from the first pass. Sensor events which were not accompanied by a passing automobile were labeled false positive and passing automobiles unaccompanied by a sensor event were labeled false negative. These error statistics were binned by the IQM of the raw sensor pixels at the time of the event as close (<1 meter), borderline (1-1.3 meters), and far (1.3-2 meters) for later evaluation. Anything outside the 2 meter range was excluded from evaluation due to the context of the problem (greater than the width of road lane) and VL53L8's technical limit (increased error above 2 meters). Automobiles which were greater than 1 full lane of separation in the video were marked as true negative for this reason.

For the purposes of evaluating close pass detection, we treated both close and borderline passes as positive predictions and produced a separate evaluation metric for far passes. By this definition, all users experienced at least one close pass during their ride. This allows us to evaluate the performance of the system within the range of interest of 0 - 1.3 meters while still reporting the evaluation metrics for the full range of the sensors. Interestingly, during post-ride questioning, 3 out of 7 users reported experiencing at least one close pass while they actually only experienced borderlines close passes at 1-1.3 meters. This may indicate that the currently accepted legal definition of a close pass may still be too small compared to the tolerance even experienced riders are willing to accept when describing a safe passing distance. The F1-score for close pass detection was found to be 0.915 and the F1-score for far passes was found to be 0.868. These results are illustrated in Figure 9.

Of a total of 269 sensor triggers events within the range of interest, we saw only 5 triggers which were not an automobile. While already small, this can be further mitigated through aggregation in a crowdsourcing setting [19]. Additionally, due to the directness of the sensing modality, i.e, physically measuring proximity, these errors can can sometimes be interpreted given context. Qualitative examination of these errors revealed our errors to be primarily at the maximum boundary of the sensor range (close to 2 meters).

[3] While this approach could be scaled up by using object detection models for larger evaluations, we opted for a fully manual approach to ensure the highest quality labels in this initial validation.

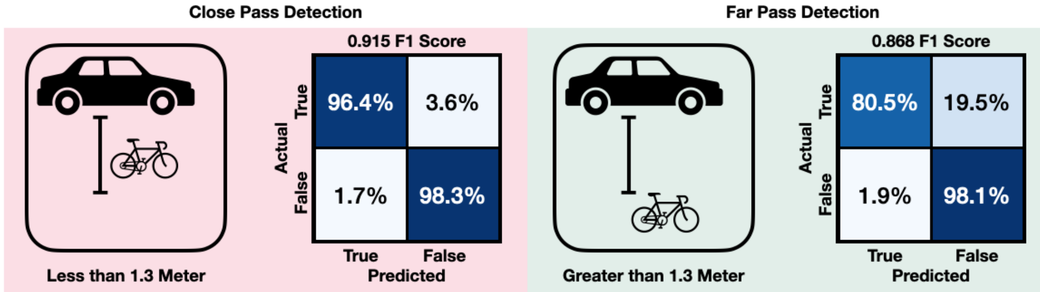

Figure 9: The classification results from detected passing objects for both the close pass and borderline close pass region (sub 1.3 meters) and the far pass region (between 1.3 and 2 meters) as well as the categories of common sources of error for both close and far.

Interestingly, our processing heuristics were able to correctly filter out most pedestrians, other cyclists, bollards, e-scooters, trees, and even motorcycles traveling both with and opposite the cyclist. This is desirable as these road users are often within the close passing margin but do not pose the same level of risk as passenger automobiles. We found 2 false positives due to static objects to the left of the cyclist while riding on footpaths, 1 false positives due to bollards on protected bicycle lanes, 1 due to a passing cyclist passing very close, and 1 due to the a participant making a very sharp turn causing the ground to trigger the sensors. In all cases, these errors occurred in scenarios which may be identified by including additional sensors. For example, using geographic context from GPS to identify riding on bollard protected bike lane or foot path where cycling near pedestrians and bollards is most likely to occur, and using the smartphone Inertial Measurement Unit (IMU) to filter out data during sharp turns.

6 Longitudinal Deployment Study

With the sensing system validated, we moved on to study the relationship of close passes with existing measures of safety across the road network through a 2 month longitudinal study with 15 cyclists. During this study participants rode as usual but with ProxiCycle installed on their bicycle to crowdsense a city-scale map of close passes which we then used to compare against two existing measures of cyclist safety: (1) a map of historic bicycle collisions provided by the Seattle department of transportation as an empirical measure of safety, and (2) a map of perceived safety which we collected through surveying the general population in Seattle.

Unlike in Section 5.2 where we recruited from the general population of cyclists, we prioritized recruiting cyclists who ride more frequently and on highly varied routes in our deployment to maximize coverage across the city. So, we recruited through a local cycling activist community: *Seattle Neighborhood Greenways* via their monthly email newsletter. As part of our screening survey, we asked participants how often they cycled (i.e., daily, multiple times a week, once a week, monthly, or more) and for the consistency of the routes they rode (i.e., on a scale of 1 to 10, with 1 being as varied as possible and 10 being same route every time). After 2 weeks we received 206 responses from potential participants. More than half of our respondents had drop bars on their primary bicycle and about a third owned an Android phone. We built 15 sensor modules and recruited from the subset of drop bar and android users to standardize both our application and physical sensor design. We sorted respondents in order of self-reported ride frequency and route variability and recruited from the top of this list working down. This recruitment process resulted in one participant from the study in Section 5.2 participating in the longitudinal deployment as well. Participants were not compensated and participated purely through volunteer interest.

Each participant received a single ProxiCycle device and a brief instruction on how to install the device on their bicycle, connect over bluetooth to their phone, and interact with the paired application to disconnect and confirm data sharing. Data was periodically uploaded to a remote server. Over the next 2 months we recorded 2050 close passes over 1604 kilometers (1002 miles) traveled over 240 unique bike rides. The map in Figure 10 A) shows a heat map of the density of close and borderline passes recorded across the geographic area covered by these rides. At the end of the deployment, we correlated the locations of these close passes with existing measures of cyclist safety in a spatial analysis.

6.1 Spatial Analysis

To understand the relationship between close passes and existing measures of cyclist safety, we correlated the locations of close passes (Figure 10 A) with both a survey of people's perceived safety at different points on the road network (Figure 10 B) and historic automobile-to-bicycle collisions (Figure 10 C). We discuss these evaluations in the following sections.

6.1.1 Close Passes as an Indicator of Collisions. Recall that collision surrogates, like our proposed measure of close passes, should be spatially collocated with historic collisions but occur more frequently to serve as an indicator of safety within a shorter temporal window [24]. To demonstrate the quality of close passes as a collision surrogate, we compare the locations of recorded close passes over our 2 month deployment as a indicator of empirical collision recorded and shared publicly by the local department of transportation in the last 5 years. These spatial signals are shown in Figure 10 A) and C), respectively. The intentions of this analysis are not to claim close passes alone serve as a comprehensive model of risk, but that close passes are correlated with existing measures of risk

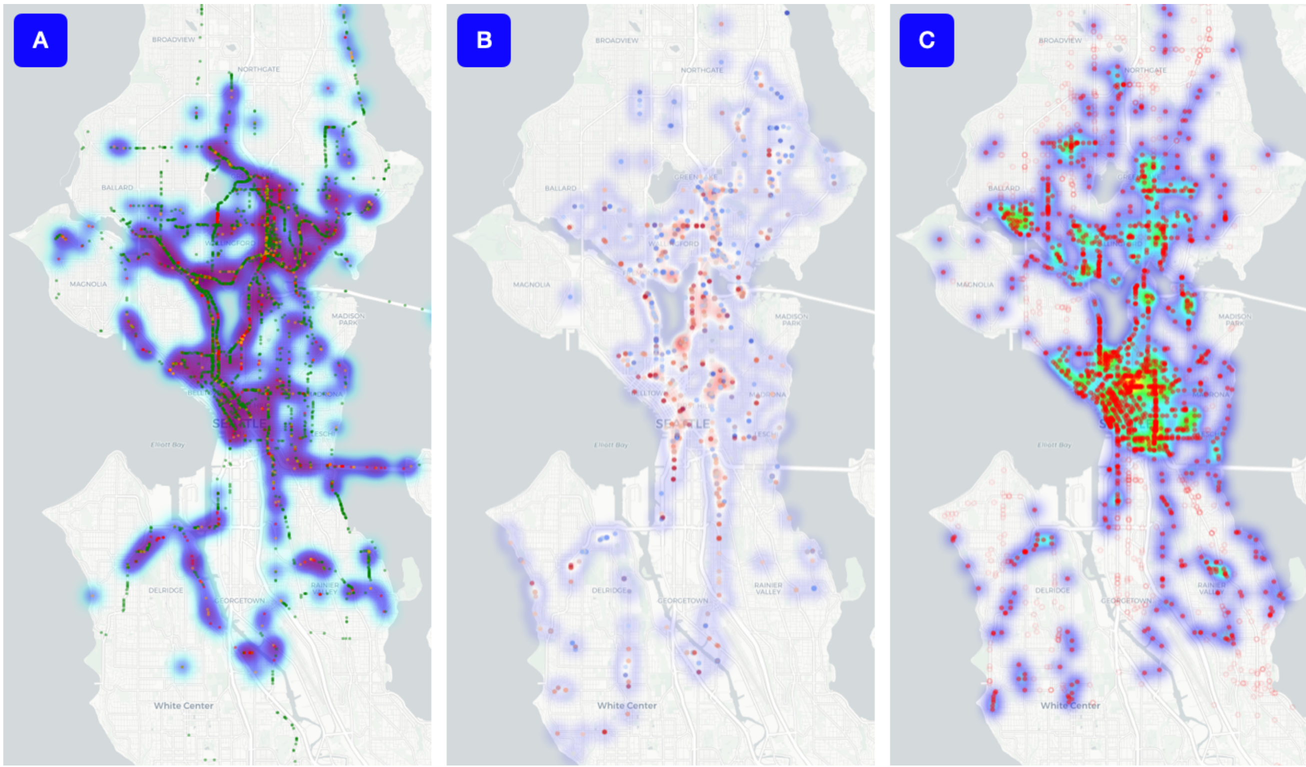

Figure 10: A) shows a heat map of close pass density across the routes traveled by participants during our study. Colored points are overlaid at each sensor event trigger including legal far passes, borderline close passes (between 1 meter and 1.3 meters), and close passes (less than 1 meter) in green, orange, and red respectively. B) Shows a heat map and points overlaid of survey respondent's perceived safety rating at locations sampled from our participant's routes. Blue points indicate higher perceived safety while red points indicate higher perceived risk. C) Shows a heat map of historic collisions including at least one bicyclist and one automobile over the last 5 years within a 200 meter radius of our participant's routes. Red points representing collision locations are overlaid, with those outside the range of our participant's route plotted as opaque.

and can be a useful signal for studying and modeling cycling risk in the absence of sufficient or current collision data.

We are only able to analyze locations where we have collected data, so we sampled locations across the routes traveled by our cyclists every 100 meters and assigned them a label of high risk if there was a historic collision within the distance margins within the last 5 years and low risk otherwise. We then use the presence of close passes within the same margin to indicate this risk. We show results of distance margins from 100 to 250 meters in Table 1, reporting the results for 200 meters inline. We chose to limit the window of historic collisions to 5 years to avoid including outdated collisions which potentially occurred prior to updated infrastructure, while maintaining a large enough window to represent safety across the entire road network (i.e, while 10 years of collisions holds more statistical power, collisions from 10 years ago may not represent the safety of the current road network). We chose to report the 200 meter distance margin as this encompasses the length of a typical block (i.e., distance between two intersections) in the location of our study. This resolution allows cyclists to plan deliberate detours to avoid collision hotspots. Finer resolutions (shorter than a block) may be too small to afford rerouting around hot spots and would inevitably be aggregated to a coarser resolution to be useful.

We found spatial Pearson correlation of close passes and historic collisions to be 0.6 with $P < 0.0001$. We found a precision-recall area under curve (PR-AUC) of 0.83 and macro-F1 score of 0.8 at estimating the locations of collisions using locations of close passes at a 200 meter distance margin. The results across all distance margins are shown in Table 1. These results show a significant correlation between the locations of close passes and collisions, indicating that close passes may be a useful collision surrogate in the absence of collision data. Additionally, close passes occur at a much higher frequency and therefore can provide a more up-to-date view of safety across the current road network in a shorter window of time (2 months vs. 5 years).

6.1.2 Close Passes as an Indicator of Perceived Safety. While less indicative of risk than collisions, people's perceptions of safety is often used in place of collision data to study safety across the road network where collision data is sparse or unavailable [65, 68]. We surveyed people's perceptions of safety at the locations with and without observed close passes in our study to compare close passes as an indicator of perceived safety.

In order to collect a measure of perceived safety to compare against, we conducted a survey study (N=76, 21 with self-reported cycling experience below 4) recruited through snowball sampling from the same activist group as our deployment participants. In the survey, participants were provided Google Street View (GSV) links which visualized locations where close passes had been observed during our 2-month deployment and were asked to rate their perception of safety with respect to traffic at each location on a 7-point Likert scale (from 1 - very safe to 7 - very unsafe). Additionally, we included an equivalent number of locations where no close passes were observed during our deployment to get a balanced sample of perceived safety across both potentially dangerous and potentially safe locations. This resulted in a set of 1027 locations to be rated which were subdivided into 27 mini-surveys containing 50 locations each (half where close passes were observed and half where they were not). Each survey participant provided basic demographic information including age, gender, and experience level cycling (prompted similarly to Section 3 that experience is defined as combination of frequency and consistency cycling), along with safety ratings for each location and whether they were familiar with the location or not. Some survey participants completed multiple mini-surveys rating more than 50 locations. All locations were rated by at least 2 raters allowing for an inter-rater agreement to be calculated. We used average of pairwise Cohen's Kappa across all pairs of raters with overlapping locations which we found to be 0.42 which is considered "fair to good" by prior studies [9]. The results of this survey are illustrated in Figure 10 B). Additionally, we found that inexperienced group (i.e., raters with experience less than or equal to 4) skews more in favor of assigning unsafe ratings to randomly surveyed locations. This follows our expectation that less experienced cyclists are less likely to feel safe and are more likely to benefit from a map of cyclist safety. This is shown in Figure 11. As a comparison, Figure 11 also shows the distributions of safety ratings assigned by the experience group for locations they were familiar or unfamiliar with. The increased overlap in the safety ratings assigned between familiarity groups indicates that familiarity with the specific location was less of a factor in influencing the participants assigned safety rating than their own experience cycling.[4]

After collecting survey responses, we compared the perception of safety at each location to the presence of close passes within a distance margin. To do this, we considered locations with an average perceived safety rating above a 4 as "high risk" and any locations less than or equal to 4 as "safe". Because of the limited agreement across survey responses and limited quantity of survey responses in general, we were unable to generate a correlation with significance of $P<0.05$. Despite this, when comparing the locations with close passes to the locations with perceived risk, we find a PR-AUC of 0.85. The results at other distance margins are available in Table 1. This indicates that locations where close passes were observed are more likely to be perceived as "high risk" on average while locations where no close passes were observed are more likely to be perceived as "safe".

[4]The experience participants were familiar with roughly half (783 of 1617) locations rated by them.

Figure 11: (Top) A Kernelized Density Estimate (KDE) showing the distribution of safety ratings assigned by inexperienced (N=21) and experienced (N=55) cyclists. Participants were determined to be experienced by a cut-off applied to self-reported experience of 4 or more on a 7-point likert scale (i.e., participants who self identify as "neutral" to "very experienced"). (Bottom) A similar KDE showing the distributions of safety ratings assigned by experienced cyclists for locations which they were and were not familiar with as a baseline.

6.1.3 Perceived Safety as an Indicator of Collisions. For completeness, we correlated our perceived safety survey responses to collisions as a baseline. This results in a PR-AUC of 0.82, indicating that close passes are similarly accurate to perceived safety surveys, but are significantly more scalable with little to no user effort and can provide continuous views into road safety as infrastructure changes. For example once ProxiCycle is installed, close passes can be sourced entirely passively without user input and continuously over time while perceived safety requires routine user reporting which does not scale. For context, our perceived safety surveys required approximately 30 minutes to complete for 50 locations, which across 76 users is almost 40 hours of labeling, while ProxiCycle requires only minutes to install. Additionally we found close passes measured using ProxiCycle to to have a higher correlation with higher significance to 5-year collision data than perceived safety survey responses.

6.2 Participant Feedback

At the end of the deployment we ask all participants to join an optional 20-30 minute exit interview. 8 of the 15 participants could attend and accepted the interview. These 8 participants accounted for nearly 80% of distance traveled during our study. Prior to the interview study we shared a custom interactive map visualizing each user's routes and associated close passes to reflect on prior to the interview. The interviews were semi-structured and open-ended, conducted remotely over Zoom while the user's map was screen shared. Each participant was asked to share their experience using the sensor as well as their thoughts on the resulting close pass hot spots (specific to their own routes) accumulated over the study. Participants were also asked whether they thought the hotpots indicated anything about their rides and if so what? We also asked them to identify who might benefit from the insight from this close pass hot spot map and why? The interviews were conducted by a single author who made notes and used Affinity Diagrams to generate themes across participants [77]. We quickly reached saturation [84] and 3 themes emerged from these interviews.

6.2.1 Theme 1: Utility to Novice Cyclists, policy makers, and planners. A common theme across participants was the utility to primarily novice cyclists and potentially policy makers. *"Absolutely, this would be very informative for people who are new to biking in an area and haven't figured out their own route preferences yet, especially in places like Seattle where it's all about trade off's like how much topography you're willing to deal with. In order to find your routes, you need to see objective data and the more data you have the better."* (P6) Another participant recalls their time as a beginner cyclist mapping the city themselves *"I would say if I hadn't ridden these routes before, and I saw a map like this, it would impact where I might choose to ride... For example, when the bike lane got put up on Roy, I literally didn't know about it so I would ride on Mercer which is really not a great place to ride. Programs like Strava where they show you where people ride is nice, but it's not clear if it's just a race or actually people riding"* (P2). This raises the importance of safety context along with empirical measures of route popularity. Some participants noted that close passes may be more indicative of rider comfort than true safety: *"I think this data would be useful because close passes don't necessarily hurt anyone but they contribute to a general feeling of unsafely. There's right hooks and dooring which may be more severe crashes but your data is capturing something that impacts more of how people choose to interact with biking"* (P5). This again highlights the importance of comfort and perceived safety on lowering the barriers to cycling adoption.

6.2.2 Theme 2: Impact of Rider and Driver Behavior on Utility. Most participants raised awareness to potential confounds due to their own behavior during data collection. For example, participants raise concerns of the lane position they ride in leading to closer or further passes. *"I was actually hit by a car on Roosevelt but since then my personal choice has been to take more quiet routes at the expense of time. My bike style is very much ride in the middle of the lane. My personal data points probably aren't the most useful, but with more data and more riders, the utility of the data might improve"* (P4). Another participant raises the opposite consideration with their data *"I am actually more biking off to the side of the lane because that's where I feel safer so that's why there are fewer close passes than expected here."* (P3). Both of these participants identified behaviors that may lead to fewer close passes compared to their expectation of an average cyclist due to their riding style, despite citing opposite factors (i.e., taking up more or less of the lane). Users also noted variation in driver behavior across time leading to different results. *"A lot of bad places, no close passes showed up. But I do feel like cars were giving me a wider berth. I don't know if something has changed at the beginning of the month or a coincidence of the time. But I was hoping the sensor would capture this and it just didn't happen even though it has happened dozens of times over the last few years"* (P1). Similarly participants noted the time of day impacting the signal. P5 and P3 raised the fact that they rode the most dangerous routes in the early morning when few cars pass, leading to the illusion of safety on this route if time of day is not considered. Finally, some participants noted the importance of driver speed to contextualize close pass data. *"Another thing is the faster I'm going, the less I care about cars passing me because they are passing me at a similar speed. New cyclists are likely going slower and so it must be much scarier for them"* (P5). Another participant notes, *"Keep in mind that downtown things generally move slower so close passes are less scary. In suburban roads which are paint only and higher speeds, behavior dynamics are different. I'm more assertive and take the whole lane here but close passes here are much worse due"* (P6). This feedback demonstrates the importance of looping in temporal and environmental context whenever using close passes for downstream analyses.

6.2.3 Theme 3: Ease of Data Acquisition Making Up For Errors. Despite multiple participants raising concerns about potentially misleading information, all participants agreed that the data was valuable and easy enough to collect to merit using the system. *"There's definitely some improvements but it's easy enough to do and I'm fairly motivated to collect data because I see It as useful"* (P5). Another notes, *"I'd absolutely keep using this. I find it pretty easy to use and I kind of treat it like my lights. I already charge things and take things off my bike when I'm not riding. It kind of fits into my routine"* (P4). Interestingly, the use of multiple bicycle mounted tools is a trait common amongst experienced cyclists, further motivating this tool as something experienced cyclists can use to record data which can be shared with novice cyclists to propagate expertise amongst the community.

7 Discussion

In this section we will discuss ideal applications for ProxiCycle along with an action plan for scalable adoption. Then we discuss the limitations of first sensing close passes and then utilizing close passes for inferring safety across the road network, along with opportunities for future work.

7.1 Application Scenarios

This work demonstrates the feasibility of deploying ProxiCycle to study cycling safety in the absence of historic safety priors and on a shorter timescale than existing approaches, enabling quicker and more continuous evaluation.

7.1.1 Routing for Novice Riders and Outreach. We envision a number of applications which this affords. First and foremost, ProxiCycle can be used to bootstrap bicycle safety studies in municipalities lacking any records of reported bicycle collisions, perceived safety, or rides. Cycling activist groups could deploy ProxiCycle with a few riders to estimate safety across the road network which can be used to advocate for infrastructure improvements. This signal can then be used in navigation applications to direct cyclists on known safer routes or can be provided as a map layer for data exploration.

	Estimating Collisions				Estimating Perceived Safety				Perceived Safety Baseline			
Distance Resolution	Corr	ROC AUC	PR AUC	F1	Corr	ROC AUC	PR AUC	F1	Corr	ROC AUC	PR AUC	F1
100m	0.40	0.71	0.63	0.69	–	0.52	0.74	0.49	0.12	0.58	0.62	0.57
150m	0.51	0.74	0.74	0.75	–	0.52	0.82	0.46	0.15	0.60	0.76	0.56
200m	**0.60**	**0.79**	**0.83**	**0.80**	–	0.53	0.85	0.46	–	0.60	0.82	0.53
250m	0.71	0.85	0.87	0.84	–	0.52	0.88	0.44	0.16	0.61	0.88	0.51

Table 1: A table showing the performance of using locations of close passes to estimate both locations of collisions and perceived safety as measures of risk. The last column presents baseline measures of using perceived safety survey responses as the estimator of collision locations. Metrics are shown at different distance margins for spatial aggregation to demonstrate performance at different resolutions. Pearson's correlation is provided for all comparisons. Correlation is omitted in comparisons with P>0.05. F1 score is calculated using Macro-F1.

These mapping tools could then be used for education and outreach to promote cycling as related work [67] and our formative study in Section 3 indicate that knowledge of safer routes can be the deciding factor for some people on whether to take the first step into cycling. ProxiCycle can also provide useful feedback to novice riders on the safety of their chosen routes to expedite the development of their mental maps.

7.1.2 Planning. Additionally, we envision ProxiCycle being useful for A/B testing navigation routes as systems built alongside ProxiCycle could suggest alternative routes and automatically measure the safety across them using the quantity of observed close passes on each route in real time. Similarly, data gathered by ProxiCycle users before and after infrastructure development can be used in empirical statistical analysis to quantify the impact of bicycle related development on bicycle safety. Empirical data like this can help support future investment into healthy infrastructure by demonstrating its immediate impacts on safety.

7.1.3 Collision Modeling. Further, since close passes occur much more frequently than collisions, they can be used in data-hungry analyses which may be infeasible with more scarce collision data such as comparing safety across time of day. Similarly, the locations of close passes could be used to train data-hungry models to estimate problematic locations or even predict future risk at scale. These models could then be used for more dynamic safety-based navigation which accounts for changes in bicycling safety over time on shorter timescales that what is possible with collision data.

7.2 Adoption Plan

Like any crowdsourcing system, ProxiCycle depends on user adoption. In Section 6 we saw the potential for grassroots neighborhood organization to bootstrap adoption. Similarly, much of the original location-aware computing work leaned heavily on hobbyist "War-Drivers" to map wireless networks. [41]. In Section 6.2 we found our participants valued contribution to sourcing close pass data worth the effort to ride with ProxiCycle. While ProxiCycle devices could be built by hobbyist or distributed by activist organizations, we also see room for engagement within the private sector through distribution from insurance companies or bike-share companies. Many organizations from governments to private companies (often in collaboration with governments) offer economic incentives for bicycle transportation from discounts at local businesses all the way up to e-bicycle purchase rebates[5]. A similar program could exist for distributing ProxiCycle stakeholders with an interest in understanding safety (i.e., governments, insurance companies, etc.) offering rebates in return for volunteering to use ProxiCycle on rides.

7.3 Limitations & Future Work

While this work demonstrates the feasibility of passively crowdsourcing close passes as a measure of road safety or comfort, there are multiple limitations and opportunities for future work. Primarily, close passes offer a view into the risk of a certain type of collision, so-called "side-swipes", but may not represent near-collisions of different formats such as "dooring" or "right-hook" collisions. Analyses utilizing close pass data may therefore benefit from additional safety context. Similarly, historic traffic levels were unavailable at the time of this analysis, but future studies could leverage traffic data to studying the weighted correlation in close pass and collision rates with respect to total trips at each location. Additionally, since the absence of a close pass does not inherently mean 0% risk, this analysis can characterize the stability of close pass rates over different time windows.

The performance of close passes as a measure of safety may also differ across road type based on bicycle infrastructure treatment, road width, and other features. A follow-up analysis could investigate the correlation of close passes with existing safety measures across road type and the presence of cycle infrastructure or other factors such as time of day or weather conditions. Similarly, additional sensors could be introduced such as low-resolution cameras or microphones to measure qualitative differences in the types of vehicles committing close passes such as size or speed to further investigate the severity of close passes at different locations, i.e., larger and faster vehicles are strictly more dangerous [82]. Finally, future studies could incorporate cyclist activity recognition to incorporate cyclist behavior context (i.e., lane position or riding speed) which may alter the performance of close passes as an effective collision surrogate.

[5]Anecdotally, one of our participants noted during our interview that their employer offers a program which funds a bicycle purchase and routine maintenance for the employee if the employee commits to commuting to work at a minimum of twice a week

8 Conclusion

In this paper, we presented ProxiCycle, a standardized and scalable approach for crowdsourcing a measure of cyclist safety along the road network by sensing the proximity of passing cars as a collision surrogate. We motivate our approach through a formative study, test its technical performance in a controlled environment, and the deployed the system in a real-world city-wide longitudinal study with 15 users for 2 months. We validated our system's technical performance in both an in-lab and real-world environment and demonstrate a high degree of accuracy both at close pass detection (<1.3 meter range) and legal pass detection (1.3-3 meter range) with F1-scores of 0.915 and 0.868 respectively. We then compare the close passes recorded by this system during real-world use against both historic bicycle collision data and user-reported perceived safety finding a significant Pearson's correlation of 0.6 in locations of close passes in just 2 month of deployment with 5 years of collision data. We show that these close passes can estimate empirical cycling risk across the road network in the absence of existing historic collision and can converge in less time. We believe this work is a valuable step in advancing a line of inquiry into bicycle safety by demonstrating the feasibility of ProxiCycle as a fully passive and easily scalable sensing platform for measuring cycling safety across the road network.

References

[1] [n. d.]. UpRide. https://upride.cc/. Accessed: 1-24-2024.

[2] The United States Environmental Protection Agency. 2021. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2021. https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2021. Accessed: 1-31-2024.

[3] Ammar Al-Taie, Graham Wilson, Euan Freeman, Frank Pollick, and Stephen Anthony Brewster. 2024. Light it Up: Evaluating Versatile Autonomous Vehicle-Cyclist External Human-Machine Interfaces. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.

[4] Apple. [n. d.]. Use Crash Detection on iPhone or Apple Watch to call for help in an accident. https://support.apple.com/en-us/104959. Accessed: 1-20-2024.

[5] Joshua S Apte and Chirag Manchanda. 2024. High-resolution urban air pollution mapping. *Science* 385, 6707 (2024), 380–385.

[6] Joshua S Apte, Kyle P Messier, Shahzad Gani, Michael Brauer, Thomas W Kirchstetter, Melissa M Lunden, Julian D Marshall, Christopher J Portier, Roel CH Vermeulen, and Steven P Hamburg. 2017. High-resolution air pollution mapping with Google street view cars: exploiting big data. *Environmental science & technology* 51, 12 (2017), 6999–7008.

[7] David Castells-Graells, Christopher Salahub, and Evangelos Pournaras. 2020. On cycling risk and discomfort: urban safety mapping and bike route recommendations. *Computing* 102 (2020), 1259–1274.

[8] Wu Chen, Trine Agervig Carstensen, Ranran Wang, Sybil Derrible, David Rojas Rueda, Mark J Nieuwenhuijsen, and Gang Liu. 2022. Historical patterns and sustainability implications of worldwide bicycle ownership and use. *Communications Earth & Environment* 3, 1 (2022), 1–9.

[9] Domenic V Cicchetti and Sara A Sparrow. 1981. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American journal of mental deficiency* 86, 2 (1981), 127–137.

[10] Bike Citizen. [n. d.]. Bike Citizen Home Page. https://www.bikecitizens.net/en. Accessed: 4-11-2024.

[11] Commonwealth of Massachusetts. 2024. Massachusetts General Laws, Chapter 90, Section 14: Precautions for Safety of Other Travelers. https://www.mass.gov/info-details/mass-general-laws-c90-ss-14 Accessed: 2024-11-09.

[12] Kirsten K Davison, Jessica L Werder, and Catherine T Lawson. 2008. Peer reviewed: Children's active commuting to school: Current knowledge and future directions. *Preventing chronic disease* 5, 3 (2008).

[13] Code de la Route. 1975. Royal Decree of 1 December 1975 establishing general regulations on road traffic policing and the use of public roads. https://www.code-de-la-route.be/fr/reglementation/1975120109/cyclistes~hra8v386pu

[14] Bundersministerium der justiz. 2024. Road Traffic Regulations (StVO) § 5 Overtaking. https://www.gesetze-im-internet.de/stvo_2013/__5.html#:~:text=%C2%A7%205%20%C3%9Cberholen,als%20der%20zu%20%C3%9Cberholende%20f%C3%A4hrt.

[15] Shane B Eisenman, Emiliano Miluzzo, Nicholas D Lane, Ronald A Peterson, Gahng-Seop Ahn, and Andrew T Campbell. 2010. BikeNet: A mobile sensing system for cyclist experience mapping. *ACM Transactions on Sensor Networks (TOSN)* 6, 1 (2010), 1–39.

[16] Nicholas N Ferenchak and Wesley E Marshall. 2020. Validation of bicycle level of traffic stress and perceived safety for children. *Transportation research record* 2674, 4 (2020), 397–406.

[17] Brian Ferris, Kari Watkins, and Alan Borning. 2010. OneBusAway: results from providing real-time arrival information for public transit. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1807–1816.

[18] Colin Ferster, Trisalyn Nelson, Karen Laberee, and Meghan Winters. 2021. Mapping bicycling exposure and safety risk using Strava Metro. *Applied geography* 127 (2021), 102388.

[19] Andrew Fox, BVK Vijaya Kumar, Jinzhu Chen, and Fan Bai. 2015. Crowdsourcing undersampled vehicular sensor data for pothole detection. In *2015 12th annual IEEE international conference on sensing, communication, and networking (SECON)*. IEEE, 515–523.

[20] Peter G Furth, Theja VVK Putta, and Paul Moser. 2018. Measuring low-stress connectivity in terms of bike-accessible jobs and potential bike-to-work trips. *Journal of transport and land use* 11, 1 (2018), 815–831.

[21] Garmin. [n. d.]. Garmin Varia. https://www.garmin.com/en-US/p/698001. Accessed: 1-24-2024.

[22] Kevin Gildea, Daniel Hall, and Ciaran Simms. 2021. Configurations of underreported cyclist-motorised vehicle and single cyclist collisions: Analysis of a self-reported survey. *Accident Analysis & Prevention* 159 (2021), 106264.

[23] Google. [n. d.]. Google Fit. https://www.google.com/fit/. Accessed: 4-11-2024.

[24] Feng Guo, Sheila G Klauer, Jonathan M Hankey, and Thomas A Dingus. 2010. Near crashes as crash surrogate for naturalistic driving studies. *Transportation Research Record* 2147, 1 (2010), 66–74.

[25] Oddrun Helen Hagen and Maja Karoline Rynning. 2021. Promoting cycling through urban planning and development: a qualitative assessment of bikeability. *Urban, planning and transport research* 9, 1 (2021), 276–305.

[26] Susan Handy, Bert Van Wee, and Maarten Kroesen. 2014. Promoting cycling for transport: research needs and challenges. *Transport reviews* 34, 1 (2014), 4–24.

[27] Kotaro Hara, Victoria Le, and Jon E. Froehlich. 2013. Combining Crowdsourcing and Google Street View to Identify Street-level Accessibility Problems. In *Proceedings of the 31st international conference on Human factors in computing systems* (Paris, France) *(CHI '13)*. 10 pages. https://doi.org/tbd

[28] Herbert William Heinrich. 1941. *Industrial Accident Prevention. A Scientific Approach.* McGraw-Hill.

[29] Amreen A Imrit, Jaimy Fischer, Timothy CY Chan, Shoshanna Saxe, and Madeleine Bonsma-Fisher. 2024. A Street-Specific Analysis of Level of Traffic Stress Trends in Strava Bicycle Ridership and its Implications for Low-Stress Bicycling Routes in Toronto. *Findings* (2024).

[30] Irene Isaksson-Hellman. 2012. A study of bicycle and passenger car collisions based on insurance claims data. In *Annals of Advances in Automotive Medicine/Annual Scientific Conference*, Vol. 56. Association for the Advancement of Automotive Medicine, 3.

[31] Peter L Jacobsen. 2015. Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury prevention* 21, 4 (2015), 271–275.

[32] Rui Jiang, Peng Wu, and Chengke Wu. 2022. Driving factors behind energy-related carbon emissions in the US road transport sector: A decomposition analysis. *International journal of environmental research and public health* 19, 4 (2022), 2321.

[33] Wenqiang Jin, Srinivasan Murali, Youngtak Cho, Huadi Zhu, Tianhao Li, Rachael Thompson Panik, Anika Rimu, Shuchisnigdha Deb, Kari E Watkins, Xu Yuan, et al. 2021. CycleGuard: A Smartphone-based Assistive Tool for Cyclist Safety Using Acoustic Ranging. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4 (2021), 163–1.

[34] Marilyn Johnson, Judith Charlton, Jennifer Oxley, and Stuart Newstead. 2010. Naturalistic cycling study: identifying risk factors for on-road commuter cyclists. In *Annals of advances in automotive medicine/annual scientific conference*, Vol. 54. Association for the Advancement of Automotive Medicine, 275.

[35] Ahmet-Serdar Karakaya, Jonathan Hasenburg, and David Bermbach. 2020. SimRa: Using crowdsourcing to identify near miss hotspots in bicycle traffic. *Pervasive and Mobile Computing* 67 (2020), 101197.

[36] Ahmet-Serdar Karakaya, Thomas Ritter, Felix Biessmann, and David Bermbach. 2023. CycleSense: Detecting near miss incidents in bicycle traffic from mobile motion sensors. *Pervasive and Mobile Computing* 91 (2023), 101779.

[37] Ali Kassim, Hasan Tayyeb, and Mubark Al-Falahi. 2020. Critical review of cyclist speed measuring techniques. *Journal of traffic and transportation engineering (English edition)* 7, 1 (2020), 98–110.

[38] Christopher E Kline, Charles H Hillman, Bonny Bloodgood Sheppard, Bethany Tennant, David E Conroy, Richard F Macko, David X Marquez, Steven J Petruzzello, Kenneth E Powell, and Kirk I Erickson. 2021. Physical activity and sleep: An updated umbrella review of the 2018 Physical Activity Guidelines Advisory Committee report. *Sleep medicine reviews* 58 (2021), 101489.

[39] Alexandra Knight and Samuel G Charlton. 2022. Protected and unprotected cycle lanes' effects on cyclists' behaviour. *Accident Analysis & Prevention* 171 (2022), 106668.
[40] Richard Kuzmyak and Jennifer Dill. [n. d.]. Walking and Bicycling in the United States. https://onlinepubs.trb.org/onlinepubs/trnews/trnews280www.pdf. ([n. d.]). Accessed: 1-21-2024.
[41] Anthony LaMarca, Yatin Chawathe, Sunny Consolvo, Jeffrey Hightower, Ian Smith, James Scott, Timothy Sohn, James Howard, Jeff Hughes, Fred Potter, et al. 2005. Place lab: Device positioning using radio beacons in the wild. In *Pervasive Computing: Third International Conference, PERVASIVE 2005, Munich, Germany, May 8-13, 2005. Proceedings 3*. Springer, 116–133.
[42] Anneka R Lawson, Vikram Pakrashi, Bidisha Ghosh, and WY Szeto. 2013. Perception of safety of cyclists in Dublin City. *Accident Analysis & Prevention* 50 (2013), 499–511.
[43] Chu Li, Katrina Ma, Michael Saugstad, Kie Fujii, Molly Delaney, Yochai Eisenberg, Delphine Labbé, Judy L Shanley, Devon Snyder, Florian P Thomas, et al. 2024. "I never realized sidewalks were a big deal": A Case Study of a Community-Driven Sidewalk Audit Using Project Sidewalk. (2024).
[44] Luyang Liu, David Racz, Kara Vaillancourt, Julie Michelman, Matt Barnes, Stefan Mellem, Paul Eastham, Bradley Green, Charles Armstrong, Rishi Bal, et al. 2023. Smartphone-based hard-braking event detection at scale for road safety services. *Transportation research part C: emerging technologies* 146 (2023), 103949.
[45] Greig Logan, Camilla Somers, Graham Baker, Hayley Connell, Stuart Gray, Paul Kelly, Emma McIntosh, Paul Welsh, Cindy M Gray, and Jason MR Gill. 2023. Benefits, risks, barriers, and facilitators to cycling: a narrative review. *Frontiers in Sports and Active Living* 5 (2023).
[46] University College London. [n. d.]. 100 Cyclists Project. https://www.ucl.ac.uk/civil-environmental-geomatic-engineering/research-projects/2022/oct/100-cyclists-project. Accessed: 1-31-2024.
[47] Kevin Lynch. 1960. *The Image of the City*. Massachusetts Institute of Technology Press, Cambridge, Massachusetts.
[48] Richard Manton, Henrike Rau, Frances Fahy, Jerome Sheahan, and Eoghan Clifford. 2016. Using mental mapping to unpack perceived cycling risk. *Accident Analysis & Prevention* 88 (2016), 138–149.
[49] Ricardo Marqués and Vicente Hernández-Herrador. 2017. On the effect of networks of cycle-tracks on the risk of cycling. The case of Seville. *Accident Analysis & Prevention* 102 (2017), 181–190.
[50] Andrii Matviienko, Swamy Ananthanarayan, Abdallah El Ali, Wilko Heuten, and Susanne Boll. 2019. NaviBike: Comparing Unimodal Navigation Cues for Child Cyclists. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.
[51] Andrii Matviienko, Mario Boot, Andreas Löcken, Bastian Pfleging, Markus Löchtefeld, Tamara Von Sawitzky, Gian-Luca Savino, Miriam Sturdee, Josh Andres, Kristy Elizabeth Boyer, et al. 2024. Learning from Cycling: Discovering Lessons Learned from CyclingHCI. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–5.
[52] Andrii Matviienko, Florian Müller, Dominik Schön, Paul Seesemann, Sebastian Günther, and Max Mühlhäuser. 2022. BikeAR: Understanding cyclists' crossing decision-making at uncontrolled intersections using Augmented Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.
[53] Elliot Mbunge. 2020. Integrating emerging technologies into COVID-19 contact tracing: Opportunities, challenges and pitfalls. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* 14, 6 (2020), 1631–1636.
[54] Fabio Miranda, Maryam Hosseini, Marcos Lage, Harish Doraiswamy, Graham Dove, and Cláudio T Silva. 2020. Urban mosaic: Visual exploration of streetscapes using large-scale image data. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.
[55] Simone Mora, Amin Anjomshoaa, Tom Benson, Fábio Duarte, and Carlo Ratti. 2019. Towards large-scale drive-by sensing with Multi-Purpose City scanner nodes. In *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*. IEEE, 743–748.
[56] United Nations. [n. d.]. Take Action for the Sustainable Development Goals. https://www.un.org/sustainabledevelopment/sustainable-development-goals/. Accessed: 1-21-2024.
[57] The Japan News. [n. d.]. Japan Skiers' Smartphones Making Unnecessary Emergency Calls. https://japannews.yomiuri.co.jp/society/general-news/20230129-87465/. Accessed: 1-20-2024.
[58] National Conference of State Legislatures. [n. d.]. Safely Passing Bicyclists Chart. https://www.ncsl.org/transportation/safely-passing-bicyclists-chart. Accessed: 1-24-2024.
[59] Paul L Olson and Michael Sivak. 1986. Perception-response time to unexpected roadway hazards. *Human factors* 28, 1 (1986), 91–96.
[60] World Health Organization. 2020. Cyclist safety: an information resource for decision-makers and practitioners. https://iris.who.int/bitstream/handle/10665/336393/9789240013698-eng.pdf?sequence=1. Accessed: 1-21-2024.
[61] Diego Pajarito, Suzanne Maas, Maria Attard, and Michael Gould. 2018. Mapping cyclists routes: involving citizens in collecting open cycling data. In *Proceedings of the AGILE*. 7.
[62] Richard A Plunz, Yijia Zhou, Maria Isabel Carrasco Vintimilla, Kathleen Mckeown, Tao Yu, Laura Uguccioni, and Maria Paola Sutto. 2019. Twitter sentiment in New York City parks as measure of well-being. *Landscape and urban planning* 189 (2019), 235–246.
[63] Martin Porcheron, Leigh Clark, Stuart Alan Nicholson, and Matt Jones. 2023. Cyclists' use of technology while on their bike. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.
[64] Daniele Quercia, Rossano Schifanella, and Luca Maria Aiello. 2014. The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city. In *Proceedings of the 25th ACM conference on Hypertext and social media*. 116–125.
[65] Alexander Rasch, Sara Moll, Griselda López, Alfredo Garcia, and Marco Dozza. 2022. Drivers' and cyclists' safety perceptions in overtaking maneuvers. *Transportation research part F: traffic psychology and behaviour* 84 (2022), 165–176.
[66] Sasank Reddy, Katie Shilton, Gleb Denisov, Christian Cenizal, Deborah Estrin, and Mani Srivastava. 2010. Biketastic: sensing and mapping for better biking. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1817–1820.
[67] William Riggs. 2019. Perception of safety and cycling behaviour on varying street typologies: Opportunities for behavioural economics and design. *Transportation research procedia* 41 (2019), 204–218.
[68] Murilo Ribeiro Rodrigues, Antônio Nélson Rodrigues da Silva, and Inaian Pignatti Teixeira. 2022. Assessing the applicability of the cyclists' Level of Traffic Stress (LTS) classification to a medium-sized city in a developing country. *Journal of Transport & Health* 24 (2022), 101321.
[69] Gustavo Romanillos, Martin Zaltz Austwick, Dick Ettema, and Joost De Kruijf. 2016. Big data and cycling. *Transport Reviews* 36, 1 (2016), 114–133.
[70] Max Roser. 2021. Causes of death globally: what do people die from? *Our World in Data* (2021). https://ourworldindata.org/causes-of-death-treemap.
[71] Duncan Rowland, Martin Flintham, Leif Oppermann, Joe Marshall, Alan Chamberlain, Boriana Koleva, Steve Benford, and Citlali Perez. 2009. Ubikequitous computing: designing interactive experiences for cyclists. In *Proceedings of the 11th international conference on human-computer interaction with mobile devices and services*. 1–11.
[72] Manaswi Saha, Mikey Saugstad, Teja Maddali, Aileen Zeng, Ryan Holland, Steven Bower, Aditya Dash, Sage Chen, Anthony Li, Kotaro Hara, and Jon E. Froehlich. 2019. Project Sidewalk: A Web-based Crowdsourcing Tool for Collecting Sidewalk Accessibility Data at Scale. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, UK). 14 pages. https://doi.org/10.1145/3290605.3300292
[73] James F Sallis, Myron F Floyd, Daniel A Rodríguez, and Brian E Saelens. 2012. Role of built environments in physical activity, obesity, and cardiovascular disease. *Circulation* 125, 5 (2012), 729–737.
[74] Lana Samara, Paul St-Aubin, Franz Loewenherz, Noah Budnick, and Luis Miranda-Moreno. 2020. Video-based network-wide surrogate safety analysis to support a proactive network screening using connected cameras: Case study in the City of Bellevue (WA) United States. In *Proceedings of the Transportation Research Board 100th Annual Meeting, Washington, DC, USA*. 9–13.
[75] Gian-Luca Savino, Tamara Von Sawitzky, Andrii Matviienko, Miriam Sturdee, Paweł W Woźniak, Markus Löchtefeld, Andrew L Kun, Andreas Riener, and Jonna Häkkilä. 2021. Cycling@ CHI: Towards a research agenda for HCI in the bike lane. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–5.
[76] Eldon Schoop, James Smith, and Bjoern Hartmann. 2018. Hindsight: enhancing spatial awareness by sonifying detected objects in real-time 360-degree video. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
[77] Raymond Scupin. 1997. The KJ method: A technique for analyzing data derived from Japanese ethnology. *Human organization* 56, 2 (1997), 233–237.
[78] Tourism Shane Ross Minister for Transport and Sport. 2019. S.I. No. 495/2019 - Road Traffic (Traffic and Parking) (Amendment) Regulations 2019. https://www.irishstatutebook.ie/eli/2019/si/495/made/en/print
[79] ST. [n. d.]. VL53L8CX. https://www.st.com/en/imaging-and-photonics-solutions/vl53l8cx.html. Accessed: 1-31-2024.
[80] Strava. [n. d.]. Strava. https://www.strava.com/register-now?utm_source=google&utm_medium=cpc&utm_campaign=RSR_B_SR_AllUser_US&utm_term=strava&utm_content=Brand_Universal&gad_source=1&gclid=Cj0KCQiAn-2tBhDVARIsAGmStVlXtHldyIcNdBpJ7T_pG40TD05cksPKO6aNxF5SwTN0cYFhjJcS90saAoJhEALw_wcB. Accessed: 2-1-2024.
[81] Yeran Sun, Yunyan Du, Yu Wang, and Liyuan Zhuang. 2017. Examining associations of environmental characteristics with recreational cycling behaviour by street-level Strava data. *International journal of environmental research and public health* 14, 6 (2017), 644.
[82] Justin Tyndall. 2021. Pedestrian deaths and large vehicles. *Economics of Transportation* 26 (2021), 100219.
[83] Velo. [n. d.]. Velo.ai. https://www.velo.ai/product/copilot. Accessed: 1-20-2024.
[84] Susan C Weller, Ben Vickers, H Russell Bernard, Alyssa M Blackburn, Stephen Borgatti, Clarence C Gravlee, and Jeffrey C Johnson. 2018. Open-ended interview

questions and saturation. *PloS one* 13, 6 (2018), e0198606.
[85] SHIMPACHI YOSHIDA. 2023. Revised traffic law aimed at making roads safer for cyclists. https://www.asahi.com/ajw/articles/15092275

A Appendix

A.1 What is your gender?

(1) Woman
(2) Man
(3) Non-Binary
(4) Prefer not to disclose
(5) prefer to self-describe (fill in the blank)

A.2 How old are you?

(1) <18
(2) 18 to 24
(3) 25 to 29
(4) 30 to 34
(5) 45 to 54
(6) 55 to 64
(7) >65

A.3 How do you identify?

(1) American Indian or Alaskan Native
(2) Asian / Pacific Islander
(3) Black or Arfican American
(4) Hispanic
(5) White / Caucasian
(6) Multiple Ethnicities / Other

A.4 How would you rate your experience level cycling from 1 (no experience cycling) to 7 (very experienced i.e. cycles regularly)?

A.5 (Optional) is there anything you would like to clarify about your answer to the previous question? For example, is all of your experience cycling from outside the US?

A.6 How dangerous do you perceive general cycling as a mode of transit (i.e. cycling on the road for commuting)? (1 = Very Safe to 7 = Very Dangerous).

A.7 (Optional) is there anything you would like to clarify about your answer to the previous question?

A.8 Please rank the factors below based on how likely they are to ENCOURAGE/INFLUENCE you to cycle as a mode of transit. It is ok to rank multiple the same. (1 = Least important, 3 = Neutral, 5 = Most important)

(1) Fitness
(2) Preferred weather conditions
(3) Safety from traffic
(4) Aesthetics of environment along route
(5) Preferred elevation of route

(6) Access to bike infrastructure (i.e. bike lanes, bike storage, etc.)
(7) Safety from crime (i.e. theft, assault, etc.)

A.9 (Optional) Are there any other factors not listed which impact your decision to cycle as a mode of transit?

A.10 Please rank the factors below based on how likely they are to DISCOURAGE/PREVENT you from cycle as a mode of transit. It is ok to rank multiple the same. (1 = Least important, 3 = Neutral, 5 = Most important)

(1) Getting sweaty
(2) Weather conditions
(3) Risk of harassment
(4) Danger due to car traffic
(5) Elevation of route
(6) Lack of infrastructure
(7) Risk of crime (i.e. theft, assault, etc.)

A.11 (Optional) Are there any other factors which are not listed which prevent you from cycling as a mode of transit?

A.12 if there was a navigation/map application which could show you quantifiable measures of cycling safety at different locations (like walkscore but for cycling safety) and could provide you with the best cycling route to minimize danger from traffic without compromising other aspects such as travel time, how likely would you be to use it? (1 = Not likely to use it, 5 = Very likely to use it)

A.13 Would this application make you more likely to choose cycling as a mode of transit or cycle more often than you already do? (1 = Not likely to cycle more often, 5 = Very likely to cycle more often).

A.14 (Optional) Is there anything else you'd like to share? Feel free to share anecdotes or other factors which influence your decision to cycle or not to cycle.