RESEARCH WHITE PAPER

Anticipating the Crash

How AI Will Close the Gap in Road Saftey

GEOTAB®

Table of contents

Detected, Not Protected

In 2022, **7,522 pedestrians and 1,105 cyclists were killed on U.S. roads**, the highest cyclist toll in decades. Cameras, sensors and AI models that can warn drivers already exist, but they share a critical weakness: the datasets used to train them contain cyclists in only **about 6% of labeled instances.**

Even when these systems spot a cyclist, they mostly stop at detection. They can report "bicycle in frame", but they typically cannot assess the level or risk of a collision developing, or how much time remains.

A research team at Queen's University set out to close that gap, supported by Geotab, the City of Kingston and the Natural Sciences and Engineering Research Council of Canada (NSERC). Their work produced a first-of-its-kind dataset, new prediction models and methods that allow safety systems to perform reliably across changing environments.

From "Cyclist Detected" to "Collision Likely"

Some existing datasets captured cyclist detection; others recorded crashes, but none were built to answer the question "is this cyclist about to be hit?". The team compared **11 datasets spanning the relevant capability dimensions**, and none combined cyclist-specific dashcam video with the full range of labeling needed for collision prediction.

So they built one; CycleCrash comprises 3,000 dashcam video clips drawn from publicly available urban driving footage, split between a range of dangerous scenarios and safe cycling.

Different annotators labeled each clip three times, reconciling results across 14 attributes covering collision dynamics, cyclist behavior and scene context. The dashcam perspective is deliberate: it mirrors the view from cameras already installed in fleet vehicles.

CycleCrash supports five prediction tasks that go beyond detection. The first two assess the current situation: how risky is this cyclist's behavior and who has the right of way? The next two look forward: will a collision occur and, if so, when? The fifth estimates severity. When combined, they form the chain a collision warning system needs: from "there is a cyclist" to "a collision is likely, here is how bad it will be and here is how much time you have."

↑3K DASHCAM CLIPS

3,000 dashcam video clips drawn from publicly available driving footage, split between a range of dangerous and safe cycling.

↓11 DATASETS SURVEYED

11 existing datasets were surveyed. None maintained cyclist-specific dashcam video and labeling needed for collision prediction.

↑20% MORE ACCURATE

VidNeXt was about 20% more accurate than the next-best method.

To benchmark these tasks, **the team built VidNeXt**, a video prediction model designed to capture the way a situation unfolds over time by processing sequences of dashcam frames rather than analyzing each one in isolation.

VidNeXt was the strongest model overall, leading in half the evaluation criteria across the five tasks. On collision severity, it was about 20% more accurate than the next-best method. It also showed being the most accurate at determining who has the right of way in cyclist-vehicle interactions. No prior benchmarks existed for cyclist collision prediction, so these results establish the first reference points.

In practical terms, a system built on this research could take the same dashcam feed already running in commercial vehicles and assess how dangerous a situation is, who should yield and how much time remains before contact.

Trained in One City, Failing in the Next

Standard pedestrian detectors usually miss from **9 to 18 percentage points** more pedestrians when deployed outside their training environment. For a fleet operating across regions, a system that meets safety benchmarks at headquarters could fall dangerously short on an unfamiliar route.

Part of the problem is architectural. Conventional detectors traditionally rely on assumptions about expected object sizes. A delivery truck fits the template easily. A child on a bicycle or a wheelchair user does not. **ObjectBox, developed by the same team**, takes a different approach: it locates objects by their center point and determines boundaries from that reference, without preset size assumptions. The approach was selected for presentation at a leading computer vision conference in 2022, a distinction given to fewer than one third of submitted papers.

6%
TRAINING DATA

Cyclists are ~6% of training data. Conventional detectors rely on expected object sizes, missing smaller vehicles such as bycicles.

9-18%
POINTS MISSED

Pedestrian detectors miss from 9-18 percentage points more pedestrians when deployed outside their training environment.

↑18%
POINTS INCREASED

Mixing synthetic data into training sets improved detection by up to 18.5 percentage points at matched resolutions.

Architecture alone does not explain the full gap. Collecting and labeling real-world training images is expensive, and so a separate line of work used AI image generation to produce realistic pedestrian images with automatic labels. Mixing this synthetic data into training sets improved detection by up to **18.5 percentage points** at matched resolutions, with smaller but consistent gains at production resolution. The result is a practical way to strengthen detection without the cost of labeling thousands of real-world images.

Learning Rare Categories Improved 67% of Common Ones

When a deployed safety system is updated with data from a new environment, it tends to forget what it previously learned. This is especially dangerous for rare categories. A model re-trained on data dominated by cars and adult pedestrians, for example, can quietly lose its ability to detect children, wheelchair users and other less common road users. The system does not fail visibly - it simply stops seeing the people who need protection most.

To address this, the team developed **continual learning methods that protect critical knowledge** during updates. On a benchmark of real-world images with natural class imbalance, their approach achieved the best results among all methods tested.

The results held a surprise: **In 67% of common-object categories, re-training to recognize rare categories actually improved performance on common ones as well. The model did not trade accuracy on cars and trucks for accuracy on children and wheelchair users. Instead, it got better at both..** For fleet operators, this means a safety system can learn from new data without forgetting what it already knows.

Predicting the Path, Not Just the Position

Detection tells you where someone is, but safety requires knowing where they will be. A pedestrian at a crosswalk might continue walking, stop or reverse direction. A cyclist approaching an intersection might hold their line, merge into traffic or brake suddenly. A reliable warning system cannot bet on a single prediction - it must account for every plausible outcome.

Their trajectory prediction work accounts for all of them. Published in the leading academic journal IEEE Transactions on Pattern Analysis and Machine Intelligence, it generates the full range of plausible paths, each ranked by likelihood. If any probable path crosses the vehicle's route, that triggers a warning, even if the most likely one appears safe.

The same group developed methods that predict how a person's posture changes over time. A cyclist leaning into a turn, a pedestrian shifting weight before stepping off a curb: these are early signals that a path is about to change. Together, trajectory and posture prediction provide what current systems lack: advance notice that a road user is about to alter course. For a vehicle approaching an intersection, the difference is concrete: an alert that says "pedestrian nearby" versus one that says "pedestrian is likely to cross your path."

Built on the Footage Fleets Already Capture

Every method described here works with standard dashcam video, the same format already streaming from commercial fleet cameras - no specialized sensors, no lidar, and no proprietary data formats. The models read ordinary forward-facing footage and extract collision risk, trajectory forecasts and severity estimates.

Deploying these capabilities in real time would require compute-capable camera hardware, but the underlying data is already being collected at scale. The continual learning methods are built for systems that stay deployed and keep improving as they encounter new environments, not systems trained once and then frozen.

The Bottom Line

1. Cyclists are about 6% of training data.

CycleCrash, a 3,000-clip dataset triple-annotated across 14 labels, fills that gap.

2. AI can now predict collisions, not just detect objects.

VidNeXt was about 20% more accurate on severity than the next-best method and set the first benchmarks for cyclist collision prediction.

3. Systems trained in one city miss up to 18 percentage points more pedestrians in another.

New detection methods and synthetic training data narrow that gap without expensive manual data collection.

4. Every method works with standard dashcam video.

No specialized sensors, lidar or proprietary data formats are required.

For dashcam developers like Geotab, this research delivers a complete toolkit: a training dataset built for cyclist collision scenarios, a model architecture tested against established baselines and continual learning methods designed to keep systems improving after deployment. Every component was built around the standard video format that commercial fleet cameras already produce.

For the research community, CycleCrash fills a structural absence. Cyclists have historically appeared in only about 6% of existing training data, leaving safety AI systematically underprepared for one of the most vulnerable groups on the road. This purpose-built dataset, triple-annotated across 14 attributes and designed for public availability, now gives researchers a foundation to build, test and extend these capabilities without starting from scratch.

The dataset, the models and the methods now exist–validated, built on footage already captured by fleet cameras and available to those ready to act on them.

This white paper is based on research conducted by the Urban Scene Analytics for Road Safety (USARS) project at Queen's University, led by Prof. Michael Greenspan and Prof. Ali Etemad (Ingenuity Labs). The research was supported by Geotab Inc., the City of Kingston and the Natural Sciences and Engineering Research Council of Canada (NSERC).

Sources

National Highway Traffic Safety Administration. **Overview of Motor Vehicle Traffic Crashes in 2022**. NHTSA, 2024.

Desai, N. P., Etemad, A., & Greenspan, M. **CycleCrash: A Dataset of Bicycle Collision Videos for Collision Prediction and Analysis.** WACV 2025.

Molahasani, M., Etemad, A., & Greenspan, M. **Continual Learning for Out-of-Distribution Pedestrian Detection.** IEEE ICIP 2023.

Zand, M., Etemad, A., & Greenspan, M. **ObjectBox: From Centers to Boxes for Anchor-Free Object Detection**. ECCV 2022.

Farley, A., Zand, M., & Greenspan, M. **Diffusion Dataset Generation: Towards Closing the Sim2Real Gap for Pedestrian Detection.** CRV 2023.

Molahasani, M., Greenspan, M., & Etemad, A. **On the Relationship Between Continual Learning and Long-Tailed Recognition**. arXiv:2306.13275, 2023.

Zand, M., Etemad, A., & Greenspan, M. **Flow-based Spatio-Temporal Structured Prediction of Motion Dynamics**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023.

GEOTAB®

geotab.com