An Automatic Method to Extract Events of Drivers Overtaking Cyclists from Trajectory Data Captured by Drones

H. Vasanth Munnamgi*, Fred Feng#

*Department of Industrial and Manufacturing Systems Engineering, University of Michigan Dearborn
4901 Evergreen Rd, Dearborn, MI 48128, USA
email: munnamgi@umich.edu

Department of Industrial and Manufacturing Systems Engineering, University of Michigan Dearborn
4901 Evergreen Rd, Dearborn, MI 48128, USA
email: fredfeng@umich.edu

Keywords: cycling safety, road safety, overtaking, vulnerable road users.

1 INTRODUCTION

Cycling as a mode of transportation has been recording an upward trend in both the U.S. and Europe [1]. Unfortunately, the safety of cyclists has been a point of growing concern. Data from the National Highway Traffic Safety Administration (NHTSA) show that the crashes that occur during the events of motorists overtaking cyclists was one of the leading categories involving cyclists in fatal crashes [2]. In support of the efforts to understand the driving behavior of drivers of motorized vehicles while overtaking cyclists, this research project is aimed at developing an algorithm to identify the overtaking events.

Most existing quantitative studies on cycling safety leverage instrumented bicycles or vehicles with sensors for extracting naturalistic driving trajectories. Whereas we use data from a recent research [3] that provides naturalistic driving trajectories of road users collected at select intersections in urban areas in Germany using drones equipped with cameras. Using these videos with a data frequency of 25 Hz, the authors of this study have output inD dataset [4]. The inD dataset contains trajectories of road users that are captured in form of coordinates on a two-dimensional plane obtained from the ariel or bird's eye view of the road. Additionally, the data also captures velocity, acceleration, heading angles, dimensions of driver's vehicle etc.,

Overtaking can be thought of as four phases of approaching, steering away, passing, and returning [5]. Using the inD dataset, we have developed an algorithm to identify events when a driver of motor vehicle overtakes a cyclist. This work fits into our broader goal to contribute to the body of knowledge for improving road safety of cyclists. The work is expected to provide inputs to governmental/ traffic authorities in aspects such as design of intersections and design of bicycle lanes by providing insights into overtaking events.

2 METHODOLOGY

The goal of this paper is to develop an algorithm to automatically extract events of drivers overtaking cyclists from road user trajectory data captured by drones. Figure 1 shows one such overtaking event where the trajectories of driver and cyclist are shown in red and orange curves. The position of driver's vehicle (coded in red rectangle) and that of a cyclist (coded in orange rectangle) are plotted once every 1 second time interval. The triangles within the rectangle are intended to show the heading direction of these road users. The yellow rectangles show the event when the driver is exactly adjacent to the cyclist when in the passing phase. The algorithm provides a method to automatically extract (1) such overtaking events, and (2) safety surrogate measures such as the passing distance from the continuous trajectory data. This would allow researchers to automatically process the data to study the characteristics of drivers-overtaking-cyclists.

First, the existing time-series trajectory data of road users were studied manually to make a note of all cases where cyclists got overtook by drivers, as described in Data Labeling (sec 2.1). Second step was to deploy an algorithm that uses pre-defined thresholds for distance, angles and heading directions to identify overtaking events, (sec 2.2). Then the performance of the algorithm and the misclassifications were studied case-by-case (sec 2.3) followed by measures to tuning the algorithm (sec 2.4) by trying different values of input parameters.

2.1 Data Labeling

The algorithm proposed is aimed at classifying events of overtaking. To test the algorithm for its efficacy, the first step was to create a labeled dataset where events of overtaking were noted. Using the visualization tool [6] provided by the researchers that created the inD dataset, the data, that consists of 33 video recordings of lengths 20 to 22 minutes each, was visually studied to manually identify the overtaking events. 168 events of overtaking were identified. The output of this step was a list of unique identifiers of drivers and cyclists involved in an overtaking-event and the unique identifier of the video recording to which these road users belonged.

2.2 Developing Algorithm

An algorithm to identify the overtaking events was developed using the following variables. The algorithm identifies an event as overtaking if all the conditions of the set values are satisfied simultaneously between any of the driver and cyclist pairs.

(a) Minimum distance between vehicle and cyclist (*d*): This is obtained by computing the minimum of the shortest distance between each of sides of the cyclist and each of the sides of the driver's vehicle. Consequently, this is equal to measuring the perpendicular distance from the corner one of the road users to the side of the other road user, depending on the heading direction and the position of these users as shown in Figure 2. The condition to be satisfied for the variable is $d \leq 3.5$ m.

(b) Passing angle at driver and at cyclist: The passing angle at driver (β) is defined as the angle between the line segment joining the front-center and centroid of the driver and the line segment joining the centroids of the driver's vehicle and cyclist. The passing angle at cyclist (γ) is defined as the angle between the line segment joining the front-center and centroid of the cyclist and the line segment joining the centroids of the driver's vehicle and cyclist. These angles are shown in Figure 2. Condition to be satisfied for these variables are: $85^o \leq \beta \leq 95^o$ and $85^o \leq \gamma \leq 95^o$

(d) Heading difference (δ): Heading difference is the absolute value of difference between heading angles of driver and cyclist as shown in Figure 2. The condition to be satisfied for the variable is $\delta \leq 45^o$

(e) Speed of driver (V_m) and cyclist (V_c) respectively and a threshold speed for driver $v = 5$ km/h such that $V_m \geq v$ and $V_m > V_c$ is another condition to be satisfied. Since an overtaking could be dangerous at high speeds, any events where the driver's speed was lesser than average walking speed of $v = 5$ km/h were ignored

Figure 1: Example of driver-overtaking-cyclist.

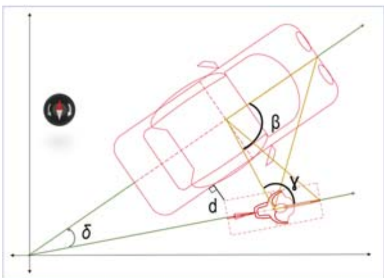

Figure 2: Passing angles and distance, heading difference.

Figure 3: Google Street View of the uniD dataset location

2.3 Performance of Algorithm – Initial Run

Out of the 168 labeled cases of overtaking events, the algorithm identified 142 (true positives), failed to identify 26 events (false negatives) and wrongly classified 6 events (false positives) as overtaking, yielding a sensitivity of 84.52% and precision of 95.95%. This was obtained for $V_m \geq v$ and $V_m > V_c$ where $v = 5$ km/h, $\delta \leq 45^o$, $d \leq 3.5$ m, $85^o \leq \beta \leq 95^o$ and $85^o \leq \gamma \leq 95^o$.

2.4 Tuning Algorithm

Investigating the false negatives, case by case, it was understood that a large portion was contributed by cases where the driver and cyclist were at the initial or final frames in the videos. Such cases needed tuning with regards to input parameters. Thus, a range of inputs were tested with β and γ in range [70º,110º] with increments of 2º, δ in range [25º,65º] with increments of 10º and d in range [2.5 m,5.0 m] with increments of 0.5 m. The classification rates were compared and the passing angles β and γ were further fine-tuned to identify the most favourable outcome.

Our algorithm identifies an event as overtaking when all these conditions are satisfied: $V_m \geq v$ and $V_m > V_c$ where v = 5 km/h, $\delta \leq 25^o$, d ≤ 3.5 m, $76^o \leq \beta \leq 101^o$ and $78^o \leq \gamma \leq 107^o$. Out of the 168 labeled cases of overtaking events, our algorithm identified 154 (true positives), failed to identify 14 events (false negatives) and wrongly classified 24 events (false positives) as overtaking, yielding a sensitivity of 91.67% and precision of 86.52%.

3 OBSERVATIONS

Out of 154 events of overtaking in the inD dataset that were identified by our algorithm, there were 18 cases or 11.69% cases of close-pass events where the passing distance was less than or equal to 1m. The passing distance had a mean = 1.48 m, median = 1.41 m, std = 0.46 m, min = 0.55 m, 25 percentile = 1.19 m, 75 percentile = 1.74 m, max = 3.44 m. There was no correlation between the passing distance and the driver passing speed, probably because of restricted traffic conditions at the intersection with slow moving traffic.

Our algorithm was applied to another dataset called uniD dataset [7] which had the trajectories recorded at a different location, RWTH Aachen University Campus. The Google Street View of a portion of the road is shown in Figure 3. The algorithm labeled 144 cases as overtaking. A manual investigation of these events revealed (a) 94 cases of overtaking events where the cyclist was on the road and was in fact overtook by driver (b) 40 cases where the cyclist was on the sidewalks/bicycle lanes (c) 9 cases where the cyclist was transitioning between the road and sidewalk/bicycle lane, almost the same level as the road (d) 1 no-overtaking.

The scatterplot of passing distance and driver speed and the histograms of the passing distance and driver passing speed of the overtaking cases that were identified by the algorithm are presented in the in Figure 2.

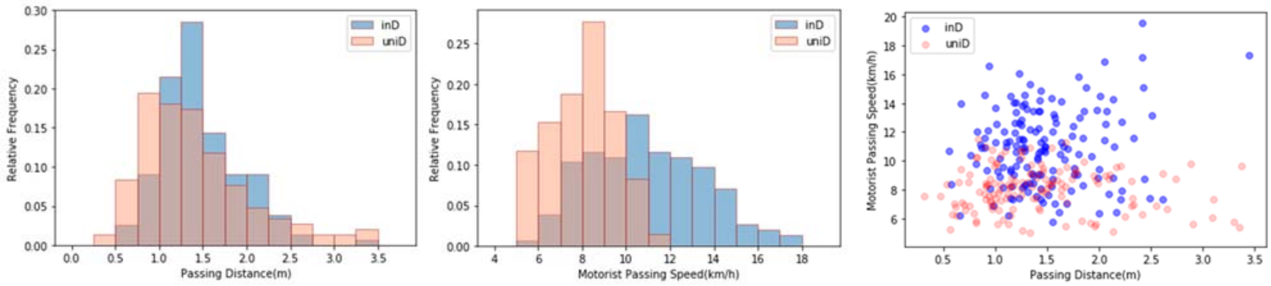

Figure 2: inD and uniD dataset: Distributions of passing distance (left) & driver passing speed and their scatterplot.

4 STUDY LIMITATIONS AND CONCLUSIONS

We developed a method that automatically extracts (1) events of drivers overtaking cyclists and (2) passing distance from trajectory data captured by drones. This method yields a sensitivity between approximately 86% and 92%. Future scope of this work is to enhance the performance of the algorithm by geo-fencing the sidewalks to reduce false positive rates. Extracting other safety surrogate measures for driver-cyclist interactions will be another extension of this work. This work will also be extended to and improved by using drone data collected at more and a wider range of locations.

REFERENCES

[1] J. Pucher and R. Buehler, “Cycling towards a more sustainable transport future,” *Transport Reviews*, vol. 37, no. 6, pp. 689–694, 2017.

[2] *Fatality Analysis Reporting System (FARS)*, National Highway Traffic Safety Administration (NHTSA). [Online]. Available: https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars. [Accessed: 29-Apr-2022].

[3] J. Bock, R. Krajewski, T. Moers, S. Runde, L. Vater, and L. Eckstein, “The IND dataset: A drone dataset of naturalistic road user trajectories at German intersections,” *2020 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1929–1934, 2020.

[4] *The inD Dataset*, Levelxdata. [Online]. Available: https://ind-dataset.com/. [Accessed: 29-Apr-2022].

[5] M. Dozza, R. Schindler, G. Bianchi-Piccinini, and J. Karlsson, “How do drivers overtake cyclists?,” *Accident Analysis & Prevention*, vol. 88, pp. 29–36, 2016.

[6] *Drone Dataset Tools*. GitHub. [Online]. Available: https://github.com/ika-rwth-aachen/drone-dataset-tools. [Accessed: 29-Apr-2022].

[7] *The uniD Dataset*. Levelxdata. [Online]. Available: https://unid-dataset.com/. [Accessed: 29-Apr-2022].