Drone-Based Trajectory Data for an All-Traffic-State Inclusive Freeway with Ramps

Ankit Anil Chaudhari[a,*], Martin Treiber[a], and Ostap Okhrin[a,b]

[a]Faculty of Transport and Traffic Sciences, Technische Universität Dresden, Dresden, Germany
[b] Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany
a.ankit_anil.chaudhari@tu-dresden.de, b.martin.treiber@tu-dresden.de
c.ostap.okhrin@tu-dresden.de
[*] Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 25, 2024

Keywords: Drone-based data collection, Trajectory data, Lane-changing behavior, Traffic state analysis, Traffic behavior analysis

1 Introduction

Traffic flow modeling plays a crucial role in transportation engineering, urban planning, and the development of autonomous vehicle technologies. Accurate modeling of driving behavior, including car-following and lane-changing dynamics, is essential for enhancing the realism and effectiveness of traffic simulation software (Li *et al.*, 2020). However, obtaining high-quality traffic trajectory data that captures the diverse range of driving behaviors across different traffic conditions has remained a challenge.

Traditional data collection methods, like stationary cameras on buildings or bridges, have limitations in providing detailed trajectory data due to restricted views and coverage, particularly in complex urban environments. Existing datasets, such as NGSIM (US Department of Transportation, 2008), often lack resolution and have numerous recording artifacts (Thiemann *et al.*, 2008, Coifman & Li, 2017).

To address these challenges, we turn to drone technology for data collection, which offers advantages like top-down views over large areas, flexibility in deployment, and ease of operation. In this study, we focus on a section of the highway in Milan, Italy, one of Europe's largest and most densely populated cities. This location provides an ideal setting to study traffic behavior in a real-world urban environment.

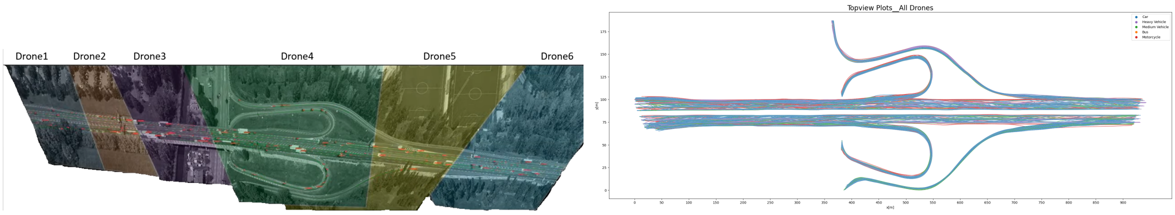
While previous research utilizing drone-based data collection has shown some promise, existing datasets (Barmpounakis & Geroliminis, 2020, Krajewski *et al.*, 2018, Ma *et al.*, 2022, Moers *et al.*, 2022) often lack the spatial and temporal coverage needed for a comprehensive analysis of driving behavior across various traffic conditions or have recorded city traffic as the Pneuma dataset collected in Athens Barmpounakis & Geroliminis (2020). While the latter is very valuable for analyzing city traffic, it suffered interruptions due to non-recorded traffic signals at intersections limiting the continuity of vehicle trajectories. Similarly, the HighD dataset (Krajewski *et al.*, 2018) and ExiD dataset (Moers *et al.*, 2022), recorded using a single drone, primarily captured free-flowing traffic conditions over a section of 420m, thus offering limited insights into driver behavior in congested traffic scenarios. The MAGIC dataset (Ma *et al.*, 2022) captured data in Shanghai using six drones. However, the data is available for each drone separately rather than stitched together preventing the analysis of trajectories of sufficient length covering several traffic stages.

In contrast, our approach leverages drone technology to overcome these limitations, enabling continuous and comprehensive data collection across different traffic states. By focusing on a high-traffic highway section in Milan, we aim to provide valuable insights into driving behavior and traffic flow dynamics, including on-ramp/off-ramp interactions, under real-world conditions.

2 Data Description and Analysis

2.1 Data Collection

The traffic data were recorded using six drones equipped with 4k resolution cameras flying at a height of 120m in a line over the A50 urban freeway in Milan, Italy, on Thursday, 27th April 2023. The recording operation started at 3:15 pm and lasted until 7:00 pm to capture the different traffic states from free flow in the afternoon to evening peak hour congested traffic. Due to the limited battery capacity of the drones, the recording time is limited to 15 minutes per flight. Nine flight campaigns, including all six drones, have been undertaken, resulting in a total recording time of 135 minutes at 30 fps and covering a 900-meter stretch of the freeway. Figure 1 (a) shows the road section and coverage under each drone.



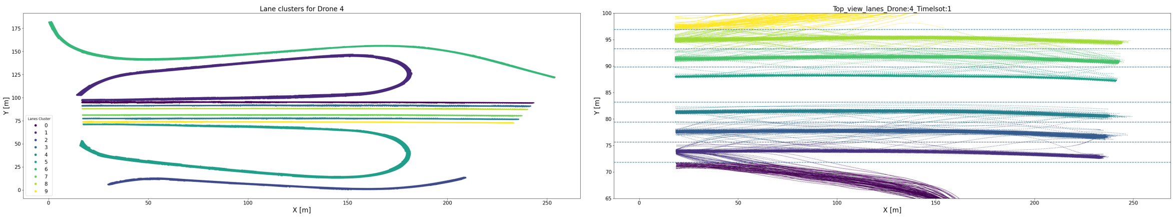
(a) *Study section coverage under six drones* (b) *Extracted Trajectory Data*

Figure 1 – *Lanes identification using clustering*

From the recorded footage, we extracted over 100,000 vehicle trajectories using single drone videos (54 datasets due to nine flight campaigns of six drones) and over 24,000 trajectories through the stitched footage of all six drones as shown in Figure 1 (b). We distinguish five categories: Cars (73%), Medium Vehicles (13.4%), Heavy Vehicles (11.3%), Buses (0.2%), and Motorcycles (2.1%). Among the total vehicles, 76.5% moved straight on the highway, while 9.7% merged and 13.8% diverged from the highway via on-ramps and off-ramps, respectively, in both directions of traffic. The integrity and precision of the extracted data were meticulously verified. Detected outliers or anomalies were removed based on stringent plausibility criteria, including continuity and physical limits.

2.2 Lane Identification

Lane identification is an important step in analyzing traffic trajectory data, as it provides valuable insights into lane occupancy and lane-changing behavior. To accomplish this, we used the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ester *et al.*, 1996) on the position data. After applying DBSCAN, clusters of data points corresponding to individual lanes were identified. Subsequently, each cluster underwent polynomial fitting to determine the best-fit line, which served as the lane center. By using these lane centers of all lanes, we were able to delineate the straight boundaries of each lane. Figure 2 (a) illustrates the identified clusters representing different lanes for the data from drone 4 by color coding, while figure 2 (b) shows the corresponding lane boundaries of the main freeway.

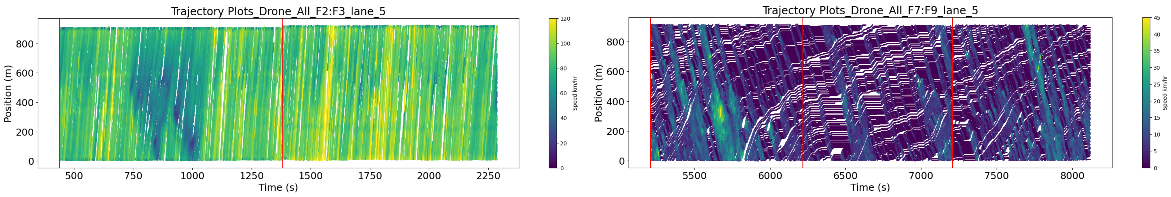


(a) *Lane clusters for Drone 4* (b) *Lane boundaries for freeway from Drone 4*

Figure 2 – *Lanes identification using clustering*

After successfully identifying the lanes, we proceeded to detect lane-changing events for each vehicle. Our analysis revealed that 46.5% of the vehicles within the dataset executed lane changes. Among these, 31.7% underwent a single lane change, while 14.8% changed lanes multiple times. Upon further examination, we observed variations in lane-changing behavior across different vehicle types. Notably, 46.5% of cars, 46.3% of medium vehicles, and 41.8% of heavy vehicles engaged in lane changes, whereas a higher proportion, 73.6%, of motorcycles, exhibited this behavior. We note that, in contrast to the city traffic in Athens (Barmpounakis & Geroliminis, 2020), the motorcycle drivers showed lane discipline.

2.3 Time-Space Plot



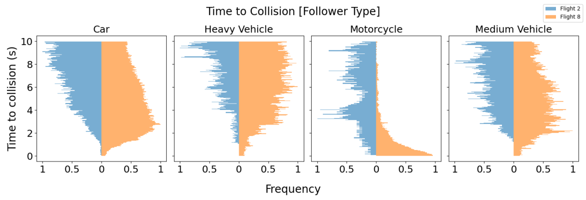
(a) *Flight Campaign 2-3* (b) *Flight Campaign 7-9*

Figure 3 – *Time-Space plots for combined trajectories of all drones*

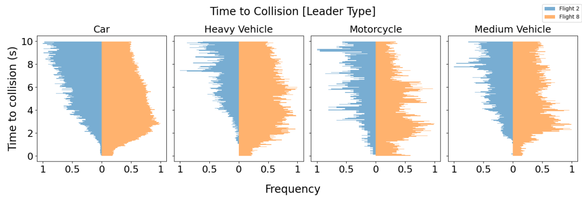
Figure 3(a) displays the trajectories of a selected lane for flight campaigns 2 and 3 (3:30 pm to 4:10 pm with an interruption in between) and showcases traffic dynamics during free-flow conditions characterized by vehicles operating at high speeds. Conversely, plot 3(b) (flights 7-9 between 5:45 pm and 6:55 pm with two interruptions) illustrates congested traffic with vehicles experiencing stop-and-go conditions. This shows the diverse range of traffic states captured in the collected data, spanning from free-flowing conditions to congested scenarios.

2.4 Time to collision

Figure 4 (a) illustrates the distribution of time to collision (TTC) for vehicles in a single lane (lane 5) under contrasting traffic conditions: free flow (flight 2) and congestion (flight 8). The distribution confirms that vehicles maintain shorter gaps between each other in congested traffic compared to free-flow conditions. Additionally, motorcycles exhibit lower gap dynamics relative to other vehicle types across all traffic conditions. In Figure 4 (b), the time to collision is analyzed in relation to different types of lead vehicles. This analysis reveals that, in free-flow conditions, the type of lead vehicle does not significantly affect the TTC while in congested traffic, motorcycles, both as a leader and as a follower, show a significantly different distribution (particularly for very low TTCs), possibly because some motorcyclists give up their lane discipline.



(a) *TTC for different types of followers*



(b) *TTC for different types of leaders*

Figure 4 – *Time to Collision for 2-flight campaigns, free-flow, and congested traffic.*

3 Discussion

Our new high-quality dataset from Milan covers different types of vehicles and shows everything from free-flow traffic to bumper-to-bumper jams. We found that almost half of the vehicles change lanes, and we also studied how traffic merges and splits at ramps; in the research, we will evaluate other descriptive statistics, such as flow-density diagrams for various lanes and sections. We expect this data to be very valuable for investigating lane changes, merges and diverges along with the longitudinal dynamics on a freeway. Furthermore, our data can help understand how drivers of different vehicle classes behave in different traffic situations. Finally, because of its high precision, the dataset can be used in developing various applications of ITS, including self-driving cars. For this purpose, we will make both the raw drone videos and the extracted trajectory data publicly available as open-source resources.

4 Acknwoldegmenet

The work is funded by DFG (Deutsche Forschungsgemeinschaft) under project grant no. 456691906, "Enhancing Traffic Flow Understanding by Two-Dimensional Microscopic Models - ETF2D".

References

Barmpounakis, Emmanouil, & Geroliminis, Nikolas. 2020. On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment. *Transportation research part C: emerging technologies*, **111**, 50–71.

Coifman, Benjamin, & Li, Lizhe. 2017. A critical evaluation of the Next Generation Simulation (NGSIM) vehicle trajectory dataset. *Transportation Research Part B: Methodological*, **105**, 362–377.

Ester, Martin, Kriegel, Hans-Peter, Sander, Jörg, Xu, Xiaowei, *et al.* 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *Pages 226–231 of: kdd*, vol. 96.

Krajewski, Robert, Bock, Julian, Kloeker, Laurent, & Eckstein, Lutz. 2018. The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. *Pages 2118–2125 of: 2018 21st international conference on intelligent transportation systems (ITSC)*. IEEE.

Li, Li, Jiang, Rui, He, Zhengbing, Chen, Xiqun Michael, & Zhou, Xuesong. 2020. Trajectory data-based traffic flow studies: A revisit. *Transportation Research Part C: Emerging Technologies*, **114**, 225–240.

Ma, Wanjing, Zhong, Hao, Wang, Ling, Jiang, Linzhi, & Abdel-Aty, Mohamed. 2022. MAGIC dataset: Multiple conditions unmanned aerial vehicle group-based high-fidelity comprehensive vehicle trajectory dataset. *Transportation research record*, **2676**(5), 793–805.

Moers, Tobias, Vater, Lennart, Krajewski, Robert, Bock, Julian, Zlocki, Adrian, & Eckstein, Lutz. 2022. The exid dataset: A real-world trajectory dataset of highly interactive highway scenarios in germany. *Pages 958–964 of: 2022 IEEE Intelligent Vehicles Symposium (IV)*. IEEE.

Thiemann, Christian, Treiber, Martin, & Kesting, Arne. 2008. Estimating acceleration and lane-changing dynamics from next generation simulation trajectory data. *Transportation Research Record*, **2088**(1), 90–101.

US Department of Transportation, FHAW. 2008. *NGSIM—Next Generation Simulation.*