What Can Public Traffic Cameras Reveal? A Short Horizon Privacy Audit Using Open Source Vehicle Tracking

Md Shadab Alam
m.s.alam@tue.nl
Eindhoven University of Technology
Eindhoven, The Netherlands

Pavlo Bazilinskyy
p.bazilinskyy@tue.nl
Eindhoven University of Technology
Eindhoven, The Netherlands

Abstract

Public traffic cameras are often treated as low risk because they rarely expose clear faces or licence plates. We study a different leakage channel: the behavioural structure extracted from ordinary public traffic footage using open source computer vision. Our preliminary pipeline converts timestamped public YouTube livestream footage from one traffic camera, spanning a 319.64 hour wall clock period, into route based vehicle events and measures how coarse signatures affect confusability. In the current run, 877,704 raw tracks were filtered into 135,488 vehicle events with full wall clock alignment. Using only class, route, and half hour time bin produced 231 baseline signatures and 178 low confusability events. Adding approximate size, speed, duration, and coarse colour increased the signature space to 11,466 signatures and 14,382 low confusability events, with 5,160 rare recurrence candidate signatures. The analysis does not perform person tracking, licence plate recognition, or exact vehicle re identification. Instead, weak, non explicit cues can make public traffic events more distinctive than aggregate counts suggest.

Keywords

Privacy, Public traffic cameras, Urban sensing, Vehicle tracking, Confusability, Behavioural privacy, Ubiquitous computing

ACM Reference Format:
Md Shadab Alam and Pavlo Bazilinskyy. 2026. What Can Public Traffic Cameras Reveal? A Short Horizon Privacy Audit Using Open Source Vehicle Tracking. In *Proceedings of Proceedings of the 2026 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2026 ACM International Symposium on Wearable Computers Adjunct (UbiComp/ISWC Adjunct '26)*. ACM, New York, NY, USA, 5 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Traffic cameras are part of everyday urban infrastructure, supporting traffic monitoring, public information, road safety, incident response, and transport planning. In intelligent transport systems, video based methods are commonly used to detect vehicles, estimate traffic flow, monitor congestion, and analyse road use [6, 12, 16]. These uses are often presented as operational rather than personal, because traffic cameras are fixed, road facing, and mainly intended to describe traffic conditions rather than individuals. At the same time, privacy regulation and public debate increasingly recognise that visual data can affect individual privacy. The General Data Protection Regulation (GDPR), European guidance on video devices, and consumer privacy regulation all reflect growing concern about automated visual sensing and personal data processing [10, 11, 22]. This creates a tension because some traffic camera footage remains publicly accessible through web pages, traffic portals, map services, or livestreaming platforms, where it can be watched, recorded, replayed, or processed by actors other than the original camera operator.

The privacy risk of such footage is not limited to whether a clear face or licence plate appears in a single frame. In high resolution public traffic streams, direct identifiers such as licence plates may sometimes be visible, which is already a data protection concern. However, focusing only on direct identifiers can understate the issue. A traffic camera also captures when vehicles appear, where they enter and leave the scene, how long they remain visible, and how they move through the road layout. These observations can become structured records even without face recognition, licence plate recognition, or exact vehicle re identification.

We refer to one completed tracked movement through the camera scene as a vehicle event. It can be described using coarse visual, temporal, and behavioural attributes such as vehicle class, route, time bin, approximate size, duration, motion, and coarse colour. We call such a combination a coarse vehicle signature. A common signature is highly confusable because many events share it, while a rare signature has low confusability because only a few events share it. Rare signatures appearing across more than one day may form candidate recurrence patterns. These are not confirmed matches of the same physical vehicle, but show how public footage can become easier to search, narrow down, and compare.

Recent traffic video research shows how ordinary road footage is increasingly treated as data rather than only as visual material. Large scale driving and urban scene datasets support scene understanding, detection, tracking, segmentation, and autonomous driving perception [7, 15, 23, 27], while generative and synthetic scene methods support domain translation, video synthesis, and controlled road scene generation [4, 14, 21, 24, 29]. Related work has also used traffic video for pedestrian analysis, traffic condition understanding, and traffic scenario generation [1, 2]. These directions motivate asking what privacy relevant structure can be extracted from publicly accessible traffic footage using accessible computer vision tools.

UbiComp/ISWC Adjunct '26, TBD

ACM ISBN 978-1-4503-XXXX-X/2026/10
https://doi.org/XXXXXXX.XXXXXXX

Object detection and tracking make this extraction practical. Detectors such as YOLO can locate vehicles in video frames [20], while tracking methods can connect detections over time to form short trajectories [5, 25, 28]. In transport research, these tools are often used for counting, flow estimation, and traffic analysis [3, 6]. The same technical process can also support privacy analysis because it turns video into repeatable searchable event records.

This concern relates to broader privacy research showing that data can remain sensitive even when direct identifiers are absent. Mobility traces can be highly distinctive from only a small number of spatiotemporal points [9], and anonymisation work has shown that removing obvious identifiers does not necessarily remove inference risk [17, 18]. Traffic camera footage has a different structure from mobile phone traces or tabular datasets, but raises a related issue because repeated observations of movement through a scene may create patterns that are distinctive even when identity is uncertain [19].

1.1 Aim of the study

This study examines short horizon privacy leakage from publicly accessible traffic camera footage using free and open source tools. Rather than using face recognition, licence plate recognition, or exact vehicle re identification, it derives route based vehicle events from public footage, represents them as coarse vehicle signatures, and measures their confusability, distinctiveness, and candidate recurrence. The aim is to show that public traffic footage can become searchable and comparable as structured behavioural data, even without identity claims that the data cannot support.

2 Method

2.1 Data source and pipeline overview

We used a public YouTube livestream of a single traffic camera at the intersection of Friant Road and Shepherd Avenue in Fresno, California, US [13]. The stream was titled *Live from Fresno, California - Supercar Spotting - Traffic Camera - Police Scanner Radio*. Its description states that the camera shows the intersection next to Woodward Park and provides real time traffic coverage intended to raise awareness about safety concerns at the intersection. The approximate centre of the intersection is 36°54'40.5"N, 119°44'33.5"W.

The analysed recording covered 16 November 2025, 07:46:52 PST (UTC−08:00) to 29 November 2025, 15:25:14 PST (UTC−08:00), corresponding to 319.64 hours of saved footage. The footage was used only for scientific analysis of aggregate and coarse vehicle event patterns, consistent with YouTube's fair use guidance that lists research as one possible fair use context [26]. The raw video was not redistributed.

The study used a two stage pipeline, shown in Figure 2. The first stage converted the downloaded public traffic camera footage into structured vehicle events, and the second used those events for privacy analysis. The pipeline did not use face recognition, licence plate recognition, or exact vehicle re identification, although the public availability of the stream means that the same footage could be processed by others using off the shelf tools. Instead, this study deliberately focuses on a weaker setting and extracts only coarse visual, temporal, and movement based signals to analyse how they form more or less distinctive vehicle event signatures.

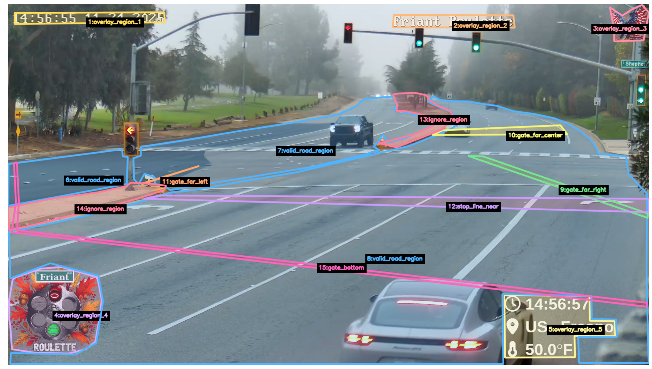

Figure 1: Scene configuration used for event extraction. The valid road region defines the analysed road area. The entry boundary captures vehicles entering the scene, while the three exit boundaries define left, straight, and right route classes. Ignore regions and overlay regions mark areas that should not be treated as vehicle movement.

2.2 Video and scene preparation

The downloaded livestream recording was standardised into 30 minute processing clips to use consistent processing settings and balance temporal continuity with computational robustness. This window preserves meaningful traffic variation, including changes in flow and signal cycles, while keeping YOLO tracking runs manageable, limiting memory and runtime failures, making intermediate outputs easier to inspect, and allowing failed or low quality portions to be handled at the clip level rather than rerunning the full recording.

The public feed contained a burned in timestamp overlay, which was treated as the trusted source for aligning each clip with wall clock time because the privacy analysis used time of appearance as one component of each vehicle event signature.

A representative frame was annotated in CVAT to define the road scene [8]. The annotation included a valid road region, overlay regions, ignore regions, a stop line, an entry boundary near the bottom of the image, and three exit boundaries representing left, straight, and right movements through the intersection. The scene layout is shown in Figure 1.

The CVAT annotations were exported and converted into the scene configuration used by the pipeline. Boundary annotations were stored as polygons and converted into centre lines. A vehicle track was considered to cross a boundary when the movement between two consecutive track centre points intersected a boundary centre line, making it possible to convert tracked movement into route labels.

2.3 Vehicle tracking and event extraction

Vehicle detection and tracking were performed with Ultralytics YOLO in tracking mode [20], using the YOLO11s model with ByteTrack as the tracking backend [28]. Tracking was restricted to road vehicle classes, including cars, motorcycles, buses, and trucks. Each tracked object was represented by its frame time, track identifier, vehicle class, confidence score, bounding box, and centre point,

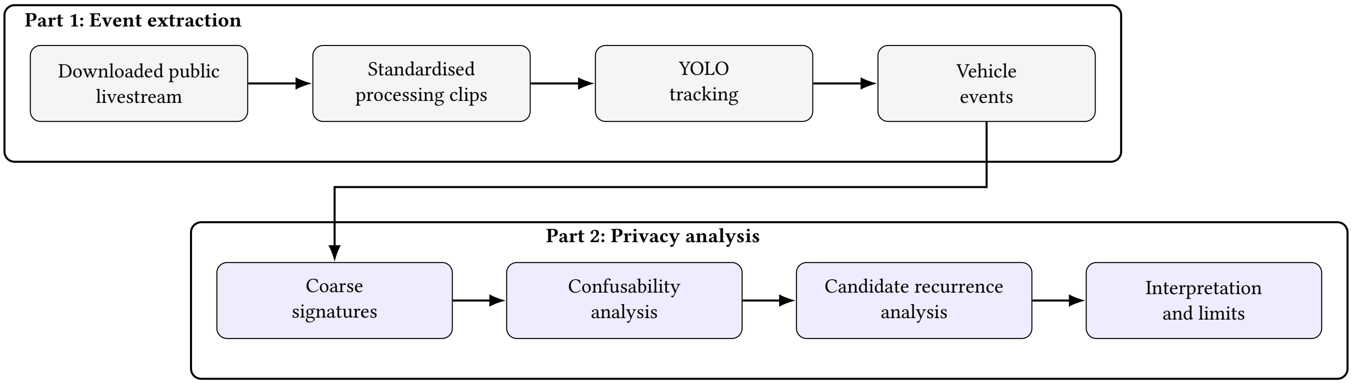

Figure 2: Pipeline overview. A downloaded public livestream from a single traffic camera is standardised into processing clips, analysed with YOLO tracking, and converted into vehicle events using scene boundaries. The extracted events are then represented as coarse vehicle signatures and analysed for confusability, candidate recurrence, and the limits of privacy claims.

which provided the basis for describing movement through the annotated road scene.

The tracking output was converted into route based vehicle events using the scene boundaries described above. A vehicle event was defined as a tracked vehicle that entered through the entry boundary and exited through one of the three route boundaries, which determined whether the event was labelled as left, straight, or right. Tracks were excluded when they had fewer than 5 track points or lasted less than 0.75 seconds, did not cross a relevant boundary, entered from an unexpected direction, or did not produce a valid route through the scene. Accepted events formed the input to the signature based privacy analysis.

2.4 Coarse signatures and privacy analysis

Privacy analysis was performed on event records rather than raw video frames. Each event was represented at three coarse signature levels. The baseline signature combined vehicle class, route type, and half hour time bin. The second level added approximate size, speed, and duration buckets. The third level added coarse colour.

The additional features were deliberately coarse. Approximate size was based on the mean area of the tracked vehicle bounding box. Motion was represented by a speed proxy, computed as the total pixel distance travelled by the track centre points divided by the observed track duration. Duration was computed from event start and end times. Size and speed were converted into four quantile based buckets, while duration used fixed thresholds of 5, 10, 20, and 40 seconds. These buckets describe broad visual and behavioural properties, not precise physical measurements.

Coarse colour was estimated from a representative frame taken from the middle of each track. The bounding box was slightly shrunk before sampling so that the estimate focused more on the vehicle body and less on nearby background. The cropped region was converted to HSV colour space and assigned to a coarse category such as dark, white, black, grey, red, orange or yellow, yellow or green, green or cyan, blue, or purple.

For each signature level, events were grouped by their signature fields, and the number of events sharing the same signature was used as the confusability count. Singleton signatures were shared by exactly one event, low confusability signatures by at most five events, and signatures above this threshold were treated as increasingly confusable as counts increased. The analysis reported unique signatures, singleton signatures, low confusability signatures, events in low confusability signatures, events per signature, and signature entropy.

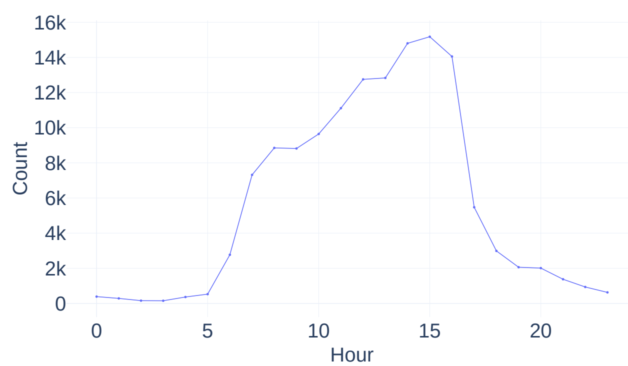

Figure 3: Hourly distribution of extracted vehicle events. The largest hourly count was 15,179 events at 15:00, and the smallest hourly count was 152 events at 03:00.

Candidate recurrence analysis was applied to rare signatures that appeared on more than one day. A recurrence candidate was defined as a signature appearing on at least two distinct days while remaining rare in the dataset, using a maximum signature count of 20 events. These candidates were not treated as confirmed vehicle matches, but only as possible recurrence patterns showing where coarse vehicle events become easier to narrow down across days. This kept the analysis focused on short horizon privacy leakage rather than deterministic vehicle re identification.

3 Results

The extraction stage considered 877,704 tracked vehicle candidates and accepted 135,488 route based vehicle events, giving an accepted

Table 1: Coarse signature summary across three feature levels.

Signature level	Unique	Singleton	Low conf.	Low conf. events	Low conf. events %	Median	Max	Recurrence
Baseline	231	39	89	178	0.13	18.00	6,046	76
Size and motion	3,800	822	1,743	3,649	2.69	7.00	724	1,555
Size, motion, and colour	11,466	4,093	7,545	14,382	10.62	3.00	537	5,160

event rate of 15.44%; all accepted events had wall clock aligned timestamps. Rejections consisted of 193,872 short tracks, 276,996 tracks with no boundary crossing, 151,929 with the wrong entry boundary, and 119,419 with an invalid exit.

The route distribution consisted of 76,233 right turn events, 55,850 straight events, and 3,405 left turn events. By vehicle class, there were 129,801 cars, 5,585 trucks, 99 buses, and 3 motorcycles. Across accepted events, the mean and median duration were 11.03 and 7.30 seconds, with a range from 1.30 to 178.00 seconds.

Figure 3 shows the hourly distribution of accepted events. The five largest hourly counts were 15,179 events at 15:00, 14,803 at 14:00, 14,055 at 16:00, 12,834 at 13:00, and 12,753 at 12:00, while the smallest hourly count was 152 events at 03:00. The five largest daily counts were 15,338 events on 21 November 2025, 14,725 on 19 November 2025, 14,383 on 18 November 2025, 14,358 on 20 November 2025, and 14,096 on 25 November 2025.

The enriched privacy feature table contained 135,485 events. Table 1 shows that adding coarse attributes substantially increased the signature space and reduced confusability. Unique signatures increased from 231 at the baseline level to 3,800 with size and motion, and 11,466 with size, motion, and colour. Low confusability signatures increased from 89 to 1,743 and then 7,545, while low confusability events increased from 178 to 3,649 and then 14,382. Candidate recurrence signatures also increased from 76 at baseline to 1,555 with size and motion, and 5,160 with size, motion, and colour.

4 Discussion

The findings indicate that the privacy implications of public traffic cameras are not limited to visible faces or licence plates. Once repeated movements through a fixed scene are converted into event records, the footage becomes searchable by time, route, vehicle class, duration, movement, size, and colour. The video is therefore no longer only a visual traffic record, but also a structured description of behavioural traces.

Scene based filtering explains why route based events are more useful for this audit than raw tracking output. Traffic footage includes occlusions, short lived detections, overlapping vehicles, edge of frame appearances, and tracks that do not correspond to complete movements through the intersection. Filtering through entry and exit boundaries focuses the analysis on completed vehicle movements rather than arbitrary detections.

The timestamp overlay and traffic imbalance help explain the privacy pattern. Time becomes part of the event signature once events are wall clock aligned, allowing observations to be grouped and compared across days. Common movements and vehicle types remain highly confusable because many observations share them, while rare movements, rare vehicle classes, or uncommon combinations of time and route provide less crowding.

The main implication is that weak signals matter when combined. Vehicle class, route, time, size, duration, motion, and colour are not strong identifiers individually, and none should be interpreted as proof of vehicle identity. However, their combination can narrow the comparison group and make some events easier to isolate. Candidate recurrence should therefore be treated as possible short horizon linkability, not as re identification.

The measures used in this audit are dataset internal. A low confusability signature means that few similar events appeared in the analysed feed, not that a vehicle is unique in the wider city or population. The audit therefore identifies where public video becomes easier to search, narrow, and compare, rather than providing an identity system.

5 Limitations and future studies

This study is based on one public traffic camera scene. Camera angle, road layout, lighting, compression, occlusion, and traffic composition all affect the extracted tracks, events, and signatures, so the results should be interpreted as a case study rather than a general estimate for all public traffic cameras. Future studies should repeat the analysis across multiple camera views, road types, cities, and traffic conditions, and characterise temporal coverage, missing periods, weather, lighting, and camera changes in the source footage.

The pipeline depends on detection and tracking quality. Missed detections, tracking switches, overlapping vehicles, night time visibility, and edge of frame appearances may affect event extraction and route labels. Future work should compare the pipeline against a manually labelled validation subset, including route correctness, class labels, event duration, and colour category reliability, to separate privacy leakage caused by the scene from artefacts introduced by the tracker.

The study does not perform exact vehicle re-identification. It does not use faces, licence plates, or direct personal identifiers, and candidate recurrence should not be interpreted as a confirmed match of the same physical vehicle. The visual and behavioural features are deliberately coarse and affected by perspective, occlusion, illumination, and compression. Future studies should evaluate privacy modified video, such as timestamp masking, lower resolution, lower frame rate, road only cropping, or background blurring, and compare privacy reduction against traffic utility.

Supplementary material

A maintained version of the code is available at https://github.com/Shaadalam9/traffic-lifelong.

References

[1] Md Shadab Alam, Olena Bazilinska, and Pavlo Bazilinskyy. 2026. A global dataset of continuous urban dashcam driving. arXiv:2604.01044 [cs.CV] doi:10.48550/arXiv.2604.01044

[2] Md Shadab Alam, Marieke H. Martens, and Pavlo Bazilinskyy. 2025. Generating Realistic Traffic Scenarios: A Deep Learning Approach Using Generative Adversarial Networks (GANs). In *Human Interaction and Emerging Technologies (IHIET-AI 2025): Artificial Intelligence and Future Applications (Open Access Science in Human Factors Engineering and Human-Centered Computing, Vol. 161)*, Taraq Ahram, Antonio Lopez Arquillos, Juan Gandarias, and Adrian Morales Casas (Eds.). AHFE International, 349–358. doi:10.54941/ahfe1005927

[3] Murat Bakirci. 2024. Utilizing YOLOv8 for enhanced traffic monitoring in intelligent transportation systems (ITS) applications. *Digital signal processing* 152 (2024), 104594. doi:10.1016/j.dsp.2024.104594

[4] Aayush Bansal, Shugao Ma, Deva Ramanan, and Yaser Sheikh. 2018. Recycle-GAN: Unsupervised Video Retargeting. In *Computer Vision – ECCV 2018 (Lecture Notes in Computer Science, Vol. 11209)*. Springer, Cham, 122–138. arXiv:1808.05174 [cs.CV] doi:10.1007/978-3-030-01228-1_8

[5] Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. 2016. Simple Online and Realtime Tracking. In *Proceedings of the IEEE International Conference on Image Processing*. 3464–3468. doi:10.1109/ICIP.2016.7533003

[6] Norbert Buch, Sergio A Velastin, and James Orwell. 2011. A review of computer vision techniques for the analysis of urban traffic. *IEEE Transactions on intelligent transportation systems* 12, 3 (2011), 920–939. doi:10.1109/TITS.2011.2119372

[7] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 3213–3223. arXiv:1604.01685 [cs.CV] doi:10.1109/CVPR.2016.350

[8] CVAT.ai Corporation. 2026. Computer Vision Annotation Tool (CVAT). doi:10.5281/zenodo.4009388 Software. Accessed: 2026-05-14.

[9] Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen, and Vincent D. Blondel. 2013. Unique in the Crowd: The privacy bounds of human mobility. *Scientific Reports* 3, 1 (2013), 1376. doi:10.1038/srep01376

[10] European Data Protection Board. 2020. Guidelines 3/2019 on processing of personal data through video devices. https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-32019-processing-personal-data-through-video_en Version 2.0, adopted 29 January 2020.

[11] European Parliament and Council of the European Union. 2016. Regulation (EU) 2016/679: General Data Protection Regulation. https://eur-lex.europa.eu/eli/reg/2016/679/oj Official Journal of the European Union, L 119, 4 May 2016.

[12] David Fernández Llorca, Antonio Hernández Martínez, and Iván García Daza. 2021. Vision-Based Vehicle Speed Estimation: A Survey. *IET Intelligent Transport Systems* 15, 8 (2021), 987–1005. doi:10.1049/itr2.12079

[13] Friant Roulette. 2025. Live from Fresno, California - Supercar Spotting - Traffic Camera - Police Scanner Radio. https://www.youtube.com/watch?v=qMYlpMsWsBE. Public YouTube livestream. Accessed: 2026-05-14.

[14] Adrien Gaidon, Qiao Wang, Yohann Cabon, and Eleonora Vig. 2016. Virtual Worlds as Proxy for Multi-Object Tracking Analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 4340–4349. arXiv:1605.06457 [cs.CV] doi:10.1109/CVPR.2016.470

[15] Jakob Geyer, Yohannes Kassahun, Mentar Mahmudi, Xavier Ricou, Rupesh Durgesh, Andrew S. Chung, Lorenz Hauswald, Viet Hoang Pham, Maximilian Mühlegg, Sebastian Dorn, Tiffany Fernandez, Martin Jänicke, Sudesh Mirashi, Chiragkumar Savani, Martin Sturm, Oleksandr Vorobiov, Martin Oelker, Sebastian Garreis, and Peter Schuberth. 2020. A2D2: Audi Autonomous Driving Dataset. arXiv:2004.06320 [cs.CV] doi:10.48550/arXiv.2004.06320 Dataset page: https://www.a2d2.audi/.

[16] Vassiliki Kastrinaki, Michalis Zervakis, and Kostas Kalaitzakis. 2003. A Survey of Video Processing Techniques for Traffic Applications. *Image and Vision Computing* 21, 4 (2003), 359–381. doi:10.1016/S0262-8856(03)00004-0

[17] Arvind Narayanan and Vitaly Shmatikov. 2008. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*. IEEE, 111–125. doi:10.1109/SP.2008.33

[18] Paul Ohm. 2010. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review* 57, 6 (2010), 1701–1777. https://ssrn.com/abstract=1450006

[19] José Ramón Padilla-López, Alexandros André Chaaraoui, and Francisco Flórez-Revuelta. 2015. Visual Privacy Protection Methods: A Survey. *Expert Systems with Applications* 42, 9 (2015), 4177–4195. doi:10.1016/j.eswa.2015.01.041

[20] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 779–788. doi:10.1109/CVPR.2016.91

[21] German Ros, Laura Sellart, Joanna Materzynska, David Vazquez, and Antonio M. Lopez. 2016. The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 3234–3243. doi:10.1109/CVPR.2016.352

[22] State of California. 2018. California Consumer Privacy Act of 2018. https://leginfo.legislature.ca.gov/faces/codes_displaySection.xhtml?lawCode=CIV&sectionNum=1798.100 California Civil Code, Section 1798.100 et seq..

[23] Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, Vijay Vasudevan, Wei Han, Jiquan Ngiam, Hang Zhao, Aleksei Timofeev, Scott Ettinger, Maxim Krivokon, Amy Gao, Aditya Joshi, Yu Zhang, Jonathon Shlens, Zhifeng Chen, and Dragomir Anguelov. 2020. Scalability in Perception for Autonomous Driving: Waymo Open Dataset. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2446–2454. arXiv:1912.04838 [cs.CV] doi:10.1109/CVPR42600.2020.00252 Dataset page: https://waymo.com/open/.

[24] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Guilin Liu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. 2018. Video-to-Video Synthesis. In *Advances in Neural Information Processing Systems*, Vol. 31. 1144–1156. arXiv:1808.06601 [cs.CV] https://papers.nips.cc/paper/7391-video-to-video-synthesis

[25] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. 2017. Simple Online and Realtime Tracking with a Deep Association Metric. In *Proceedings of the IEEE International Conference on Image Processing*. 3645–3649. doi:10.1109/ICIP.2017.8296962

[26] YouTube Help. 2026. Fair use on YouTube. https://support.google.com/youtube/answer/9783148. Accessed: 2026-05-14.

[27] Fisher Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, and Trevor Darrell. 2020. BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2633–2642. arXiv:1805.04687 [cs.CV] doi:10.1109/CVPR42600.2020.00271

[28] Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, and Xinggang Wang. 2022. ByteTrack: Multi-Object Tracking by Associating Every Detection Box. In *Computer Vision – ECCV 2022*. Springer, 1–21. doi:10.1007/978-3-031-20047-2_1

[29] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. 2017. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In *Proceedings of the IEEE International Conference on Computer Vision*. IEEE, 2242–2251. arXiv:1703.10593 [cs.CV] doi:10.1109/ICCV.2017.244