WIP: Infrastructure-Aided Defense for Autonomous Driving Systems: Opportunities and Challenges

Yunpeng Luo
UC Irvine
yunpel3@uci.edu

Ningfei Wang
UC Irvine
ningfei.wang@uci.edu

Bo Yu
PerceptIn
bo.yu@perceptin.io

Shaoshan Liu
PerceptIn
shaoshan.liu@perceptin.io

Qi Alfred Chen
UC Irvine
alfchen@uci.edu

***Abstract*—Autonomous Driving (AD) is a rapidly developing technology and its security issues have been studied by various recent research works. With the growing interest and investment in leveraging intelligent infrastructure support for practical AD, AD system may have new opportunities to defend against existing AD attacks. In this paper, we are the first t o systematically explore such a new AD security design space leveraging emerging infrastructure-side support, which we call Infrastructure-Aided Autonomous Driving Defense (I-A2D2). We first taxonomize existing AD attacks based on infrastructure-side capabilities, and then analyze potential I-A2D2 design opportunities and requirements. We further discuss the potential design challenges for these I-A2D2 design directions to be effective in practice.**

I. Introduction

As Autonomous Driving (AD) technology becomes increasingly deployed and commercialized in the real world, more and more people start to consider the security of AD vehicles. There are a lot of researches trying to create adversarial examples for fooling AI components in AD systems. Zhao et al. introduce malicious stop signs that can not be detected by AD vehicles [1], while Cao et al. create printable objects that can not be perceived by both camera and LiDAR [2], leading to serious crashes. Some other works focus on the localization module. Shen et al. use GPS spoofing t o l ead t he v ictim to crash into an incoming vehicle on the opposite lane [3].

On the other hand, the AD system design patterns are also evolving, with growing interests and investment in leveraging infrastructure-side support. Specifically, a n ew d irection of AD design called Infrastructure-Aided Autonomous Driving (IAAD) is being developed recently, which uses infrastructure side communication and sensing abilities to improve AD reliability while reducing on-board sensing cost [4]. Today, there are many ongoing IAAD testing, and even deployment efforts by companies and institutes. For example, Baidu is currently testing intersections with IAAD deployed in several cities; to demonstrate the benefits o f s uch infrastructure-side support on AD, they even showcase L4 AD capability only using infrastructure-side sensing *without using any on-boarding sensing* [5]. Seoul Robotics proposes to use sensors embedded in surrounding infrastructure to achieve L5 AD; BMW is currently testing that system at its Munich manufacturing facility [6]. HORIBA Institute for Mobility and Connectivity (HIMaC2) at UC Irvine plans to equip 25 intersections in Irvine with Velodyne's LiDAR-based intelligent infrastructure solution [7]. Among various different AD deployment scenarios today, IAAD is most attractive and thus likely to be first utilized by robo-taxi/ride-hailing services due to the cost considerations [8]. On the government side, some countries have already realized the importance and benefits of such smart transportation infrastructures not only to practical AD deployment but also to city functions (e.g., mobility and environmental aspects), and are thus making proactive policies for building such infrastructures [9, 10].

Workshop on Automotive and Autonomous Vehicle Security (AutoSec) 2022
24 April 2022, San Diego, CA, USA
ISBN 1-891562-75-4
https://dx.doi.org/10.14722/autosec.2022.23048
www.ndss-symposium.org

TABLE I: Categorization of existing AD attacks from the perspective of infrastructure-side capability requirements in I-A2D2 designs. Full version is Table II in Appendix.

Category	Attack	I-A2D2 design
A1: Perception of infrastructure-authoritative information	Sign hiding [1, 11–17]	Traffic sign
	Sign appearing [1, 16–18]	Traffic sign
	Traffic light changing [19, 20]	Traffic light
A2: Perception of dynamic road objects	Object hiding [2, 21–38]	Obstacle detection
	Object appearing [18, 20, 35, 38–42]	Obstacle detection
	Object relocation [43–46]	Obstacle detection
	Object detection disabling [35, 47]	Obstacle detection
A3: Localization	Trajectory shifting [3, 20, 48–50]	Localization
	Localization disabling [35]	Localization and lane

Considering such a new AD design trend, it is important to explore whether and how it may influence the existing AD security design space. In this paper, we are the first to systematically discuss the opportunities and challenges for such a new AD security design space leveraging emerging infrastructure-side support, which we call Infrastructure-Aided Autonomous Driving Defense (I-A2D2).

II. I-A2D2 Design Opportunities

With infrastructure-side sensing and communication abilities, many existing attacks against AD systems can potentially have new defense opportunities. Due to the different nature of the attacks, different infrastructure-side capabilities can be required for effective and systematic I-A2D2 defense designs. We thus started by performing a comprehensive survey of representative AD attacks published in recent years, and classified them into 3 categories from such I-A2D2 design requirement perspective: (**A1**) *Perception of infrastructure-authoritative information*; (**A2**) *Perception of dynamic road objects*; (**A3**) *Localization*. A short version of such categorization is in Table I; the full version is Table II in Appendix.

A. A1: Perception of Infrastructure-Authoritative Information

Opportunities. Like human drivers, AD has to follow traffic rules, such as obeying the speed limit and waiting for a red light. Quite some existing works target their perception and thus trick the victim to break traffic rules; as in Table I,

there are 8 attacks trying to hide the traffic signs from being detected by AD vehicles, e.g., using stickers [1]; 4 attacks try to create non-existing traffic signs, e.g., using a projector [17]; 3 attacks try to change the traffic light result detected by AD vehicles, e.g., by exploiting ROI designs [19].

However, since traffic signs and traffic lights are under authoritative control of the government transportation agencies, the infrastructure is able and also authoritative to provide such information. For example, the infrastructure may broadcast the existence of a stop sign at a GPS coordinate to all passing-by AD vehicles. This can at least provide another (if not more trustworthy due to the source authoritativeness) information source to AD vehicles, which thus can enable at least direct attack detection capabilities against all these existing attack vectors targeting infrastructure-authoritative information such as traffic signs and lights.

Required infrastructure-side capability. To inform the AD vehicle of authoritative information, the infrastructure must be able to communicate with the AD vehicle in time. Currently, DSRC and C-V2X are the commonly used wireless communication technologies for cars, both of which are able to deliver a message within 100ms [51, 52]. On the other hand, since the information to transmit is known in advance, the infrastructure can send it before the AD vehicle needs to react to the information, such as informing the AD vehicle 200 m away from a stop sign. Thus, if designed properly, the infrastructure should be able to always ensure the information can arrive in time to the AD vehicle side to enable effective defense design opportunities. Such authoritative information also needs to be sent with the location information, e.g., absolute GPS coordinates. The AD vehicle can then use its own real-time localization to react at the correct position, e.g., stop when it's right in front of a stop sign.

B. A2: Perception of Dynamic Road Objects

Opportunities. To avoid collision, AD vehicles need to detect dynamic road objects and avoid them proactively. Some attacks aim at the obstacle detection component and try to hide, relocate, or create non-existing objects in front of the victim. The others directly disable the obstacle detection component. We find 20 attacks trying to hide objects using a variety of methods, e.g., using malicious shape changes [2]; 8 aim at adding non-existing objects to the detection results, e.g., using sensor spoofing [35, 39, 42]; 4 try to change the location of the object in the detection results, e.g., using malicious stickers/posters [46]; and 4 try to disable obstacle detection, e.g., using sensor jamming [35, 47]. To defend against such attacks, the infrastructure-side perception capabilities (e.g., camera and LiDAR) in the emerging IAAD designs [4–7] (§I) can be leveraged to help perceive maliciously hidden objects or eliminate attacker-introduced fake objects. The AD vehicle side can fuse their own detection results with such infrastructure-side detection results, which can at least detect (if not able to correct) the attacked results.

Required infrastructure-side capability. Defending against A2 requires better communication capabilities compared to A1 in both latency and bandwidth. The AD vehicle needs to be informed of the object in time so that it can have adequate time to make decisions and react (e.g., stop before crash or change lane). Since the objects are dynamic, to achieve this the infrastructure need to frequently transmit the most recent detection results, which thus requires a much smaller latency and higher bandwidth than A1. Failing to do that, the infrastructure can end up sending obsolete information that can be even misleading in extreme cases.

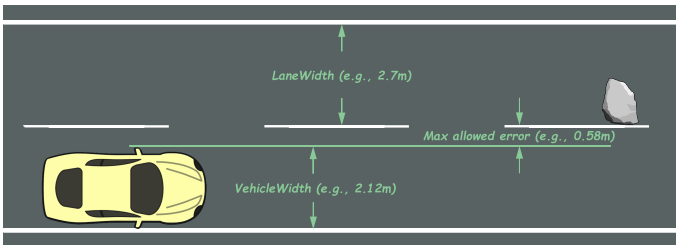


Fig. 1: Estimation of the maximum allowed obstacle perception errors for I-A2D2 designs against A2 attacks.

Besides the communication capability requirements, the infrastructure-side perception also needs to be accurate enough. To more concretely understand such accuracy requirement, we perform an estimation of the maximum allowed perception errors based on common lane width and car width. Here, the most error-sensitive scenario takes place when the detection results indicate that all obstacles are out of the current lane boundaries, but in reality the detection errors are large to the extent that there can be a misdetected obstacle with which a collision cannot be possibly avoided by the AD vehicle. Fig. 1 shows such a setup with the least error tolerance, where an obstacle is detected to locate on the boundary of the adjacent lane, but its actual location can collide with the AD vehicle even if the AD vehicle is choosing the safest possible route in the ego lane (e.g., on the rightmost side in Fig. 1). In this case, the maximum allowed detection error of such obstacle is $LaneWidth - VehicleWidth$. According to USDOT, the lane width ($LaneWidth$) of a local road should be at least 2.7 m [53]. As for vehicle width ($VehicleWidth$), we take the width of a common AD vehicle model, Lincoln MKZ, which is 2.12 m. Therefore, the maximum allowed errors for infrastructure-side dynamic object detection cannot be over 0.58 m (Fig. 1).

C. A3: Localization Attacks

Opportunities. Localization is a key component for AD to accomplish tasks such as navigation and path planning. The AD vehicles need to know where it is on the road or in the map to follow the lane and to decide whether to make a turn. It's also required to know the vehicle's ego position when using infrastructure-side authoritative information in A1 defenses (§II-A). We find 5 attacks targeting localization (Table I), e.g., using GPS spoofing [3] and dirty road patches [49], which can cause severe consequences such as driving off the road and even crashing into the incoming vehicle from the opposite direction. Since camera is usually used for lane centering in L2 AD systems, camera blinding attack [35] can also disable the AD system's localization ability, which may cause potential shifting of the vehicle position.

Similar to A2 attacks (§II-B), the infrastructure-side perception capability can enable defense opportunities against A3 attacks. Specifically, the infrastructure side can keep sending information such as locations of all perceived in-road vehicles to the AD vehicle side. On the AD vehicle side, when it first receives such information, it uses its own localization result (e.g., from GPS) to find the closest infrastructure-side detected vehicle as the representation of itself . It then performs object

tracking of such a vehicle in subsequent infrastructure-side sent frames, taking its real-time location as infrastructure-aided localization results. When the attack happens, there will be a mismatch between the AD vehicle's self-localization and such infrastructure-aided localization results; the AD vehicle can thus use the latter as an additional information source to at least perform attack detection. Note that such a design follows the trust-on-first-use (TOFU) assumption, i.e., assuming that the AD vehicle is not under localization attack the first time it receives the infrastructure-side information, so that it can correctly find the matching result from infrastructure-side detected vehicles.

Required infrastructure-side capability. Defending against A3 attacks requires similar communication capability as those against A2, as they both require real-time infrastructure-side perception. In terms of localization accuracy, in the infrastructure-aide localization discussed above, performing localization of AD vehicle is essentially performing object detection like I-A2D2 defenses against A2 (§II-B). However, comparing to A2 attacks, defending against A3 attacks needs a higher perception accuracy, as a deviation of $\frac{LaneWidth-VehicleWidth}{2}$, which is $\frac{2.7-2.12}{2} = 0.29m$ if using the same $LaneWidth$ and $VehicleWidth$ values as in §II-B, is already enough to cause a vehicle to have lane departure. To achieve such infrastructure-aided localization capabilities, it thus requires the infrastructure side to have better sensing ability and more reliable algorithm to perceive objects from the input data in the real time.

III. I-A2D2 Design Challenges

A. Precise Self-localization from Infrastructure Perception

As discussed in II-C, to defend against localization attacks (A3), the infrastructure side needs to achieve sufficiently-accurate localization of pass-by vehicles. However, we find that achieving such required accuracy is non-trivial based on our preliminary experiments in a real IAAD-deployed road.

Experiment setup. We perform our preliminary experiments on a real road where IAAD is deployed for testing purposes. The IAAD-supported road segment is over 1000 meters long with full IAAD coverage. We experiment with two LiDAR obstacle detection models on the collected data: (1) the built-in segmentation model used in Apollo 5.0 [54], an open source industry-grade AD system, which we denote as "Apollo5"; (2) PIXOR [55] from Uber ATG. Among all the objects detected in each frame, we select the one closest to the ground truth as the infrastructure-aided localization result, and compare it with the ground truth. We calculate the error in each frame to evaluate the performance.

Results. The distribution of the errors of each frame and also their median are plotted in Fig. 2. As shown, the median error of Apollo5 is 0.68 m, while that of PIXOR is 0.82 m, both of which cannot meet the requirement identified in II-C (0.29 m). To understand the causes, we examine whether the distance from the infrastructure-side LiDAR to the vehicle can affect the accuracy of localization. We collect such distance and localization errors of the Apollo5 model in each frame, and plot all data points in Fig. 3a. As shown by the red line, the errors almost monotonically increase with the distance. That is likely because when the vehicle is farther away from the LiDAR, the laser points become more sparse and thus makes

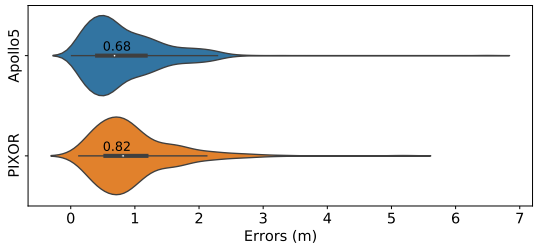
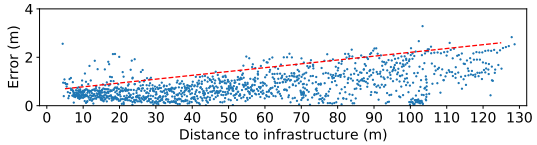
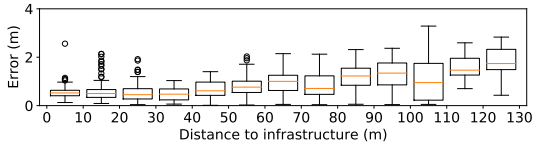


Fig. 2: Error distribution of both the Apollo LiDAR perception model ("Apollo5") and the PIXOR model.



(a) All data points



(b) Box plot

Fig. 3: Localization error of the Apollo LiDAR perception model ("Apollo5") at different distance.

it more difficult to have accurate perception. We further group data points based on distance, and draw a box plot in Fig. 3b. One interesting observation is that, when the distance is small, there are some outliers with big errors. This is likely because the LiDAR is mounted in a high position (around 3 m), and only a few laser channels with large pitching angle can cover close areas, resulting in a sparse point cloud for nearby objects.

Future improvements. We plan to explore the following directions in the future to improve this: (1) Use newer point cloud object detection models and train them with infrastructure-side data; (2) Use Kalman Filter on the infrastructure-aided localization.

B. Adaptive Attacks

Attack infrastructure-side perception. Since the sensors and the AI components used for such perception are similar to those used on AD vehicles, the attackers can apply/adapt existing vehicle-side perception attacks to the infrastructure side, or even attack both AD vehicle and infrastructure perception at the same time [2].

Exploit fixed sensor positions. Because IAAD sensors are in fixed positions, when facing the same sensor attack, infrastructure can be more vulnerable comparing to AD vehicle (e.g. no need to perform tracking and aiming for laser shooting attacks [39]). In a similar vein, generating adversarial examples can be easier as well, since it no longer requires taking the vehicle motion dynamics into consideration like in [1, 49].

New cyber-attack surface. In IAAD/I-A2D2, the communication between infrastructure and AD vehicle can expose AD vehicle's interior system to other devices, which thus introduces a new cyber-attack surface.

IV. Conclusion

In this paper, we are the first to systematically discuss the opportunities and challenges for the new Infrastructure-Aided Autonomous Driving Defense (I-A2D2) design space. We first

taxonomize existing AD attacks based on infrastructure-side capabilities, and then analyze potential I-A2D2 design opportunities and requirements. We further discuss the potential design challenges for these I-A2D2 design directions to be effective in practice. We hope that our discussions and insights can inspire more future research into this promising but currently under-explored defense design space for AD system security.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation under grants NSF CNS-1932464, NSF CNS-1929771, NSF CNS-2145493, and USDOT UTC Grant 69A3552047138.

REFERENCES

[1] Y. Zhao, H. Zhu *et al.*, "Seeing isn't believing: Towards more robust adversarial attack against real world object detectors," in *CCS*, 2019.
[2] Y. Cao, N. Wang *et al.*, "Invisible for both Camera and LiDAR: Security of Multi-Sensor Fusion based Perception in Autonomous Driving Under Physical-World Attacks," in *IEEE S&P*, 2021.
[3] J. Shen, J. Y. Won *et al.*, "Drift with Devil: Security of Multi-Sensor Fusion based Localization in High-Level Autonomous Driving under GPS Spoofing," in *USENIX Security*, 2020.
[4] S. Liu, B. Yu *et al.*, "Invited: Towards Fully Intelligent Transportation through Infrastructure-Vehicle Cooperative Autonomous Driving: Challenges and Opportunities," in *DAC*, 2021.
[5] "Baidu and Tsinghua U Introduce Apollo Air to Empower Autonomous Driving with Roadside Sensing," https://medium.com/apollo-auto/baidu-and-tsinghua-u-introduce-apollo-air-to-empower-autonomous-driving-with-roadside-sensing-99b17f50bc71.
[6] "Seoul Robotics' autonomous 'Control Tower' remotely manages self-driving vehicle fleets," https://www.engadget.com/the-level-5-control-tower-is-a-puppet-master-for-autonomous-vehicle-fleets-140041909.html.
[7] "Velodyne's LiDAR-based traffic monitoring solution to be used at 25 intersections as part of $6 million road network project in Irvine, Calif." https://www.automotiveworld.com/news-releases/velodynes-lidar-based-traffic-monitoring-solution-to-be-used-at-25-intersections-as-part-of-6-million-road-network-project-in-irvine-calif/.
[8] "To Make Self-Driving Cars Safe, We Also Need Better Roads and Infrastructure," https://hbr.org/2018/08/to-make-self-driving-cars-safe-we-also-need-better-roads-and-infrastructure.
[9] "Policy to promote autonomous driving," https://www.chinadaily.com.cn/a/202101/05/WS5ff3b702a31024ad0baa06d0.html.
[10] Y. Ge, X. Liu *et al.*, "Smart transportation in China and the United States," China Academy for Information and Communications Technology, Tech. Rep., 2017.
[11] J. Lu, H. Sibai *et al.*, "NO Need to Worry about Adversarial Examples in Object Detection in Autonomous Vehicles," *arXiv:1707.03501*, 2017.
[12] S.-T. Chen, C. Cornelius *et al.*, "Shapeshifter: Robust physical adversarial attack on faster r-cnn object detector," in *ECML PKDD*, 2018.
[13] G. Lovisotto, H. Turner *et al.*, "SLAP: Improving Physical Adversarial Examples with Short-Lived Adversarial Perturbations," in *USENIX Security*, 2021.
[14] A. Zolfi, M. Kravchik *et al.*, "The Translucent Patch: A Physical and Universal Attack on Object Detectors," in *CVPR*, 2021.
[15] R. Duan, X. Mao *et al.*, "Adversarial Laser Beam: Effective Physical-World Attack to DNNs in a Blink," in *CVPR*, 2021.
[16] D. Song, K. Eykholt *et al.*, "Physical adversarial examples for object detectors," in *WOOT*, 2018.
[17] Y. Man, M. Li *et al.*, "GhostImage: Remote Perception Attacks against Camera-based Image Classification Systems," in *RAID*, 2020.
[18] B. Nassi *et al.*, "Phantom of the ADAS: Securing Advanced Driver-Assistance Systems from Split-Second Phantom Attacks," in *CCS*, 2020.
[19] K. Tang, J. Shen *et al.*, "Fooling Perception via Location: A Case of Region-of-Interest Attacks on Traffic Light Detection in Autonomous Driving," in *AutoSec*, 2021.
[20] W. Wang, Y. Yao *et al.*, "I Can See the Light: Attacks on Autonomous Vehicles Using Invisible Lights," in *CCS*, 2021.
[21] R. Wiyatno and A. Xu, "Physical Adversarial Textures That Fool Visual Object Tracking," in *ICCV*, 2019.
[22] C. Xiao, D. Yang *et al.*, "MeshAdv: Adversarial Meshes for Visual Recognition," in *CVPR*, 2019.
[23] Y. Zhang, H. Foroosh *et al.*, "CAMOU: Learning physical vehicle camouflages to adversarially attack detectors in the wild," in *ICLR*, 2018.
[24] J. Tu, M. Ren *et al.*, "Physically Realizable Adversarial Examples for LiDAR Object Detection," in *CVPR*, 2020.
[25] Z. Wu, S.-N. Lim *et al.*, "Making an invisibility cloak: Real world adversarial attacks on object detectors," in *ECCV*, 2020.
[26] K. Xu, G. Zhang *et al.*, "Adversarial t-shirt! evading person detectors in a physical world," in *ECCV*, 2020.
[27] Z. Hau, K. T. Co *et al.*, "Object removal attacks on lidar-based 3d object detectors," *arXiv:2102.03722*, 2021.
[28] X. Zhu, X. Li *et al.*, "Fooling thermal infrared pedestrian detectors in real world using small bulbs," in *AAAI*, 2021.
[29] L. Ding, Y. Wang *et al.*, "Towards Universal Physical Attacks on Single Object Tracking," in *AAAI*, 2021.
[30] J. Tu, H. Li *et al.*, "Exploring Adversarial Robustness of Multi-Sensor Perception Systems in Self Driving," *arXiv:2101.06784*, 2021.
[31] Y. Li, C. Wen *et al.*, "Fooling lidar perception via adversarial trajectory perturbation," *arXiv:2103.15326*, 2021.
[32] J. Wang, A. Liu *et al.*, "Dual Attention Suppression Attack: Generate Adversarial Camouflage in Physical World," in *CVPR*, 2021.
[33] S. Köhler, G. Lovisotto *et al.*, "They See Me Rollin': Inherent Vulnerability of the Rolling Shutter in CMOS Image Sensors," *arXiv:2101.10011*, 2021.
[34] Y. Zhu, C. Miao *et al.*, "Can We Use Arbitrary Objects to Attack LiDAR Perception in Autonomous Driving?" in *CCS*, 2021.
[35] C. Yan, W. Xu *et al.*, "Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle," *DEFCON*, 2016.
[36] K. K. Nakka and M. Salzmann, "Indirect Local Attacks for Context-Aware Semantic Segmentation Networks," in *ECCV*, 2020.
[37] F. Nesti, G. Rossolini *et al.*, "Evaluating the Robustness of Semantic Segmentation for Autonomous Driving against Real-World Adversarial Patch Attacks," *arXiv:2108.06179*, 2021.
[38] X. Ji, Y. Cheng *et al.*, "Poltergeist: Acoustic Adversarial Machine Learning against Cameras and Computer Vision," in *IEEE S&P*, 2021.
[39] Y. Cao, C. Xiao *et al.*, "Adversarial Sensor Attack on LiDAR-based Perception in Autonomous Driving," in *CCS*, 2019.
[40] J. Sun, Y. Cao *et al.*, "Towards Robust LiDAR-based Perception in Autonomous Driving: General Black-box Adversarial Sensor Attack and Countermeasures," in *USENIX Security*, 2020.
[41] K. Yang, T. Tsai *et al.*, "Robust Roadside Physical Adversarial Attack Against Deep Learning in LiDAR Perception Modules," in *CCS*, 2021.
[42] Z. Sun, S. Balakrishnan *et al.*, "Who Is in Control? Practical Physical Layer Attack and Defense for mmWave-Based Sensing in Autonomous Vehicles," *IEEE TIFS*, 2021.
[43] Y. Jia, Y. Lu *et al.*, "Fooling Detection Alone is Not Enough: Adversarial Attack against Multiple Object Tracking," in *ICLR*, 2020.
[44] S. Jha, S. Cui *et al.*, "ML-Driven Malware that Targets AV Safety," in *DSN*, 2020.
[45] D. K. Hong, J. Kloosterman *et al.*, "AVGuardian: Detecting and Mitigating Publish-Subscribe Overprivilege for Autonomous Vehicle Systems," in *EuroS&P*, 2020.
[46] X. Chen, C. Fu *et al.*, "A Unified Multi-Scenario Attacking Network for Visual Object Tracking," in *AAAI*, 2021.
[47] D. Wang, C. Li *et al.*, "Daedalus: Breaking nonmaximum suppression in object detection via adversarial examples," *IEEE Trans. Cybern.*, 2021.
[48] Z. Kong, J. Guo *et al.*, "Physgan: Generating physical-world-resilient adversarial examples for autonomous driving," in *CVPR*, 2020.
[49] T. Sato, J. Shen *et al.*, "Dirty Road Can Attack: Security of Deep Learning based Automated Lane Centering under Physical-World Attack," in *USENIX Security*, 2021.
[50] P. Jing, Q. Tang *et al.*, "Too Good to Be Safe: Tricking Lane Detection in Autonomous Driving with Crafted Perturbations," in *USENIX Security*, 2021.
[51] Z. Xu, X. Li *et al.*, "DSRC versus 4G-LTE for Connected Vehicle Applications: A Study on Field Experiments of Vehicular Communication Performance," *Journal of Advanced Transportation*, 2017.
[52] S. Chen, J. Hu *et al.*, "A Vision of C-V2X: Technologies, Field Testing, and Challenges With Chinese Development," *IEEE IoT-J*, 2020.
[53] W. J. Stein and T. R. Neuman, "Mitigation Strategies for Design Exceptions," *US Federal Highway Administration. Office of Safety*, 2017.
[54] "ApolloAuto/apollo: An open autonomous driving platform," https://github.com/ApolloAuto/apollo.
[55] B. Yang, W. Luo *et al.*, "PIXOR: Real-time 3D Object Detection from Point Clouds," in *CVPR*, 2018.

APPENDIX

TABLE II: Categorization of existing AD attacks based on infrastructure-side capability requirements in I-A2D2 defense designs.

Category	Attack	Threat model	Impact									I-A2D2 design				
			Sign hiding	Sign appearing	Traffic light changing	Object hiding	Object appearing	Object relocation	Object detection disabling	Trajectory shifting	Localization disabling	Traffic sign	Traffic light	Lane	Obstacle location	Localization
A1: Perception of infrastructure-authoritative information	Lu et al. [11]	Stop sign poster	✓									✓				
	Chen et al. [12]	Stop sign poster	✓									✓				
	Lovisotto et al. [13]	Projection to camera	✓									✓				
	Zolfi et al. [14]	Patch to camera len	✓									✓				
	Duan et al. [15]	Laser shooting to camera	✓									✓				
	Nassi et al. [18]	Projection and billboard		✓								✓				
	Song et al. [16]	Stop sign patch	✓	✓								✓				
	Zhao et al. [1]	Stop sign patch and poster	✓	✓								✓				
	Man et al. [17]	Projection to camera	✓	✓								✓				
	Tang et al. [19]	GPS spoofing			✓								✓			
	Wang et al. [20]	IR lights			✓								✓			
	Man et al. [17]	Projection to camera			✓								✓			
A2: Perception of dynamic road objects	Wiyatno and Xu [21]	Poster				✓									✓	
	Xiao et al. [22]	Malicious object				✓									✓	
	Zhang et al. [23]	Patch on obstacles				✓									✓	
	Tu et al. [24]	Malicious object				✓									✓	
	Wu et al. [25]	Patch on obstacles				✓									✓	
	Xu et al. [26]	Adversarial T-shirts				✓									✓	
	Hau et al. [27]	Laser shooting to LiDAR				✓									✓	
	Zhu et al. [28]	Patch with bulbs				✓									✓	
	Ding et al. [29]	Malicious object				✓									✓	
	Tu et al. [30]	Malicious object				✓									✓	
	Li et al. [31]	GPS spoofing				✓									✓	
	Wang et al. [32]	Patch on obstacles				✓									✓	
	Köhler et al. [33]	Laser shooting to camera				✓									✓	
	Zhu et al. [34]	Board and drone				✓									✓	
	Cao et al. [2]	Malicious object				✓									✓	
	Yan et al. [35]	Ultrasonic foaming				✓									✓	
	Nakka and Salzmann [36]	Patch on road				✓									✓	
	Nesti et al. [37]	Billboard				✓									✓	
	Wang et al. [20]	IR lights					✓								✓	
	Cao et al. [39]	Laser shooting to LiDAR					✓								✓	
	Sun et al. [40]	Laser shooting to LiDAR					✓								✓	
	Yang et al. [41]	Malicious object					✓								✓	
	Nassi et al. [18]	Projection and billboard					✓								✓	
	Sun et al. [42]	mmWave shooting					✓								✓	
	Ji et al. [38]	Ultrasound to inertial sensors				✓	✓								✓	
	Yan et al. [35]	Ultrasound shooting				✓	✓								✓	
	Jia et al. [43]	Patch on obstacles						✓							✓	
	Jha et al. [44]	Malware						✓							✓	
	Hong et al. [45]	Compromised ROS node						✓							✓	
	Chen et al. [46]	Patch on obstacles						✓							✓	
	Wang et al. [47]	Patch							✓						✓	
	Yan et al. [35]	Laser shooting to camera							✓						✓	
	Yan et al. [35]	mmWave jamming							✓						✓	
	Yan et al. [35]	Ultrasonic jamming							✓						✓	
A3: Localization	Shen et al. [3]	GPS Spoofing								✓						✓
	Kong et al. [48]	Billboard								✓				✓		✓
	Wang et al. [20]	IR lights								✓				✓		✓
	Sato et al. [49]	Road patch								✓				✓		✓
	Jing et al. [50]	Road patch								✓				✓		✓
	Yan et al. [35]	Laser shooting to camera									✓			✓		✓