IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Survey of Navigational Perception Sensors' Security in Autonomous Vehicles

ABHIJEET SOLANKI[1*] (Graduate Student Member, IEEE), WESAM AL AMIRI[1*] (Member, IEEE), MARIM MAHMOUD [1] (Student Member, IEEE), BLAINE SWIEDER [2], SYED RAFAY HASAN[1] (Senior Member, IEEE), and TERRY N. GUO[3] (Senior Member, IEEE)

[1]Department of Electrical and Computer Engineering, Tennessee Technological University, Cookeville, TN 38505, USA.
[2]Department of Computer Science, Tennessee Technological University, Cookeville, TN, USA
[3]Center for Manufacturing Research, Tennessee Technological University, Cookeville, TN, USA
* Both authors contributed equally.

Corresponding author: Syed Rafay Hasan (e-mail: shasan@tntech.edu).

ABSTRACT The adoption of autonomous vehicles (AVs) significantly increased in recent years, playing a crucial role in enhancing transportation safety and efficiency, expanding mobility options, and reducing user costs. However, as AVs become more connected and automated, they also become more susceptible to malicious attacks, making it imperative to prioritize their security. Specifically, AVs are vulnerable to a range of cyber-threats, leading to a growing body of literature on potential cyberattacks targeting AVs. A key focus of existing research is understanding these threats and proposing mitigation strategies, particularly those aimed at attacks on navigational perception sensors, which are critical to AV navigation and decision-making. AVs rely on an array of navigational perception sensors, including LiDAR, cameras, RADAR, global positioning system (GPS) sensors, and ultrasonic sensors, to ensure secure and safe operation. However, adversarial manipulation of these sensors can compromise an AV's navigational perception system, potentially leading to catastrophic consequences. While existing surveys provide valuable insights on AV cybersecurity challenges, several gaps remain unexplored, such as hardware security issues in AVs, analyzing sensor fusion attacks and evaluating practical defense implementations. To address these gaps, this paper makes a novel contribution by offering a comprehensive review of sensor fusion vulnerabilities that have not been sufficiently addressed in previous literature, specifically focusing on the compounded effects of sensor manipulation in AVs. In addition, this paper introduces a new classification of sensor vulnerabilities, highlighting practical defense mechanisms based on real-world simulation and testbed environments, a key contribution not explored in earlier works. We also introduce a detailed classification of sensor vulnerabilities, categorize the attacks that exploit these vulnerabilities, and critically assess existing countermeasures and defense mechanisms. Our study provides unique insights into the practical implementation of these countermeasures through simulations and testbeds, which have been underrepresented in prior research. This approach aims to support the development of more robust security solutions for AVs.

INDEX TERMS Autonomous Vehicles (AVs), navigational perception sensors, Cybersecurity, CARLA simulator.

I. INTRODUCTION

THE rapid advancement of autonomous vehicle (AV) technologies has significantly improved safety, efficiency, and convenience within intelligent transportation systems (ITS), especially in light of the increasing number of accidents caused by human drivers. According to the report of National Institute of Health (NIH), human driving errors are responsible for 93% of traffic accidents [1]. Additionally, approximately 6 million car accidents occur annually, resulting in the deaths of 38,000 people [2]. Given these statistics, the widespread adaption of AVs has the potential to substantially reduce these traffic incidents, saving thousands of lives each year and enhancing overall road safety.

AVs rely heavily on perception sensors and artificial in-

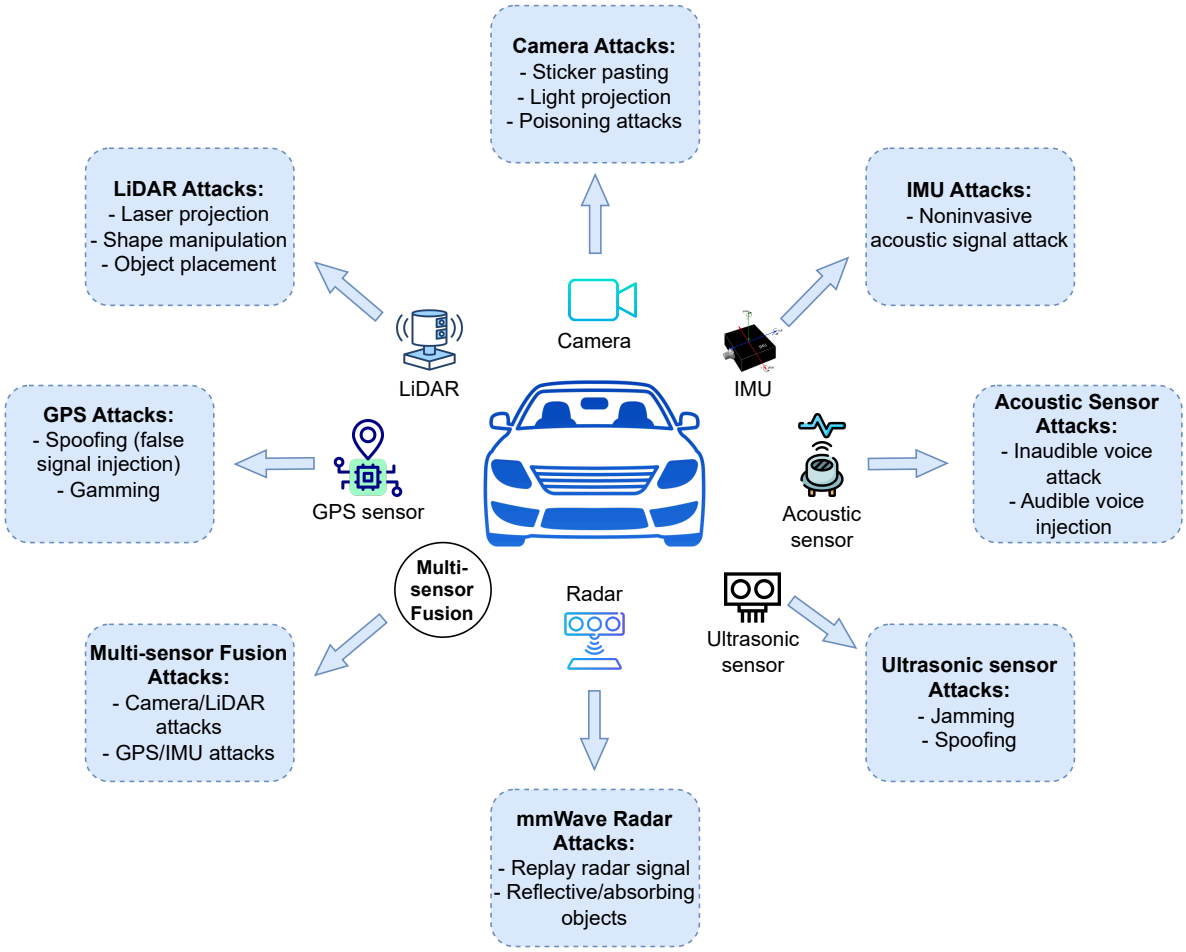


FIGURE 1. The considered security concerns on AV navigation perception sensors.

telligence (AI) to make safe and accurate navigation. AVs' navigational perception systems, integrate various sensors, including GPS for location tracking, Light Detection and Ranging (LiDAR) for creating detailed 3D maps, cameras for visual recognition, Inertial Measurement Units (IMU) for detecting movement, RADAR, acoustic sensors and cameras for object detection and distance measurement. These sensors collectively enable AVs to perceive and navigate their environment accurately, making real-time decision-making critical to their safe operation. However, the increasing reliance on these sensors has introduced a myriad of cybersecurity vulnerabilities (see Fig. 1) [3], [4]. Specifically, attacks targeting AV navigational perception sensors exploit these security weaknesses, posing severe risks to autonomous maneuvering and overall vehicle security.

Navigational perception sensor attacks involve deliberate attempts to disrupt or manipulate the sensor data used by AVs to perceive their environment for navigation. These attacks can range anywhere from GPS spoofing, which manipulates the vehicle's perceived location, to LiDAR jamming, which blinds the sensor by overwhelming it with signals. The consequences of such attacks can be severe, leading to incorrect decision-making and potentially catastrophic outcomes. As AVs become increasingly integrated into everyday life, ensuring the cybersecurity of these navigational perception systems is paramount.

The Society of Automotive Engineers (SAE) set a J3016 standard that defines six levels of driving automation, ranging from Level 0 (no automation) to Level 5 (full automation), as shown in Fig. 2 [5]. As vehicles advance through these levels, the need for accurate and secure sensor data grows exponentially. For instance, Level 3 (conditional automation) and above require the vehicle to handle most driving tasks autonomously, making the integrity of navigational perception sensor data way more crucial. At these higher levels of automation, attacks on navigational perception sensors could severely compromise the vehicle's ability to operate independently, highlighting the importance of robust cybersecurity measures.

Recently, several research papers [5]–[8] have explored

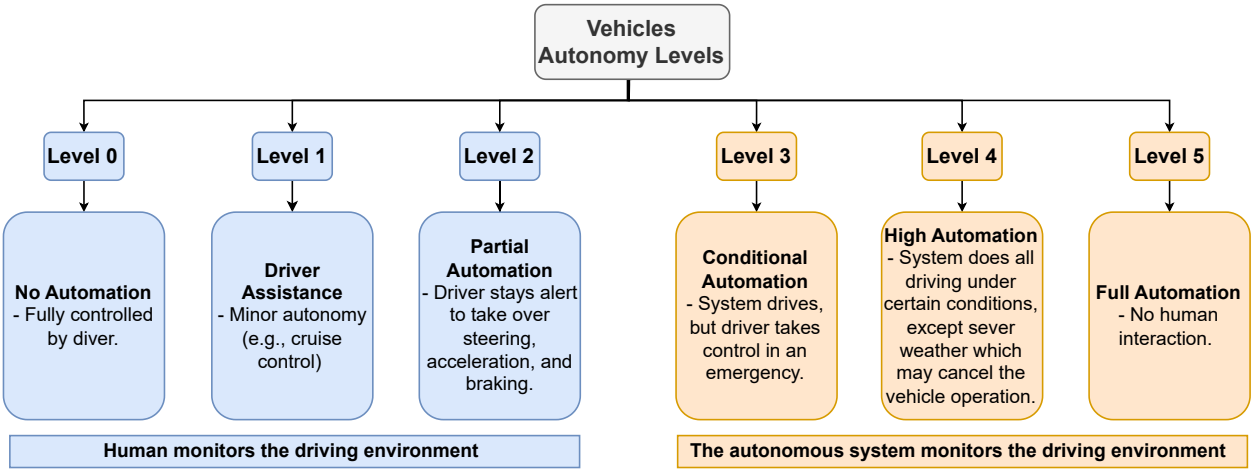


FIGURE 2. Standard autonomy levels in vehicles (adopted from [5]).

cybersecurity threats to AVs, particularly focusing on the vulnerabilities of sensors. Girdhar et. al. [5] focus on the vulnerabilities of key sensors in AVs, demonstrating contactless attack methods on a Tesla Model S and proposing countermeasures. However, the survey does not address the broader implications of sensor attacks across different AV platforms. Sharma et. al. [6] provide a wider taxonomy sensor spoofing attacks on robotic vehicles, while Niroumand et. al. [7] emphasize the role of data-driven solutions and emerging technologies. Sun et. al. [8], on the other hand, contribute by exploring digital forensics in AVs an discussing defense strategies post-incident.

While these surveys provide valuable insights, their scope is largely limited to taxonomies of attacks and high-level overviews of countermeasures. Additionally, a comprehensive and integrated analysis of navigational perception sensor security that connects theoretical threat models, concrete mitigation strategies, and experimental validation through simulation and testbed platforms.

In contrast, this paper contributes a comprehensive and application-focused survey that integrates four underaddressed but essential pillars in AV security research:

- A focused analysis of navigational perception sensor attacks (e.g., GPS, LiDAR, camera), including compound and sensor fusion-based threats.
- An up-to-date taxonomy of countermeasures, extending to off-road AV scenarios and underexplored threat conditions.
- A comparative review of simulation frameworks (e.g., CARLA and Baidu Apollo) used to emulate attacks and test defenses.
- An overview of testbed platforms for real-world experimentation and validation of sensor-related attacks and security responses.

These aspects are often only partially addressed—or entirely omitted—in earlier surveys. To the best of our knowledge, this is among the first surveys to holistically integrate simulation environments and testbed platforms alongside attack and defense taxonomies. Prior works typically focus on theoretical threat models and high-level defenses, without systematically analyzing the tools required for real-world validation. As highlighted in Table 1, no existing work provides a side-by-side comparison of attacks, defenses, simulations, and testbeds in the context of navigational sensors. Our survey aims to fill this crucial gap and provide researchers with a practical, implementable foundation for future AV security research and development.

The rest of this paper is organized as follows. Section II presents navigation perception sensor attacks. The countermeasures are presented in Section III. The AVs attacks simulation frameworks are described in Section IV. Testbed frameworks are discussed in Section V. Section VI provides key insight and lessons learned, followed by conclusions in Section VII.

Table 2 provides the major acronyms used in the paper.

II. NAVIGATIONAL PERCEPTION SENSOR ATTACKS

The vulnerabilities of navigational perception sensor attacks in AV systems arise from targeting the environment, sensors, decision-making algorithms, and control systems (see Fig. 3). Understanding the mechanics of these attacks involves defining the sequence of events from the attack source to its manifestation, known as the attack path. These paths can be categorized into specific scenarios based on their impact on different components of the AV system. As depicted in Fig. 3, there are four primary attack paths: Physical Attacks via Environment to the Sensors (PAvES) 1, PAvES 2, PAvES 3, and Deceptive Attacks via Environment to Sensors (DAvES).

PAvES attacks include scenarios where an attacker physically spoofs the sensor, leading to compromised data. For instance, PAvES 1 involves an attacker spoofing a sensor resulting in incorrect control signals being sent to the steering or throttle. PAvES 2 and PAvES 3 involve spoofing sensors to affect navigational perception data, which in turn impacts

TABLE 1. Comparison of navigational perception sensor attacks and defenses in AVs survey papers.

Survey Papers	Sensor Attacks	Defense	Limitation	Simulation	Testbed
[5]	GPS: spoofing and jamming. LiDAR: spoofing and jamming. Camera: adversarial images. IMU: tampering. RADAR: interference.	GPS: authentication and signal encryption. LiDAR: signal processing and encryption. Camera: robust algorithms. IMU: secure calibration. RADAR: frequency hopping and signal encryption.	Does not consider multi-sensor fusion and the compounded vulnerabilities that arise from sensor fusion.	Not included.	Not included.
[6]	GPS: spoofing. LiDAR: jamming. Camera: adversarial attacks. RADAR: jamming.	GPS: cross-referencing. LiDAR: redundancy and signal analysis. Camera: adversarial training. RADAR: redundancy.	Briefly describes multi-sensor fusion vulnerabilities, but lacks a detailed exploration of standardized frameworks and real-world testing environments.	Not included.	Not included.
[7]	GPS: spoofing and jamming. LiDAR: spoofing and jamming. Camera: adversarial examples, tampering. RADAR: jamming and spoofing.	GPS: multi-sensor fusion and authentication. LiDAR: signal processing and physical shielding. Camera: ML defenses and tamper detection. RADAR: signal processing and encryption.	Lacks of a detailed exploration into the practical implementation of the data-driven solutions in real-world scenarios.	Not included.	The paper discusses some testbeds that replicate real-world attacks and defense mechanisms.
[8]	GPS: spoofing and jamming. LiDAR: spoofing. Camera: adversarial attacks. RADAR: jamming.	GPS: encryption and authentication. LiDAR: secure protocols. Camera: adversarial training. RADAR: secure protocols.	Does not consider multi-sensor fusion vulnerabilities.	Not included.	Not included.
Our Paper	Focuses on navigational perception sensor attacks, with a particular emphasis on GPS, LiDAR, and cameras.	Conducts an in-depth examination of countermeasures targeting critical sensors.	Focus on perception and navigation sensors, with some emphasis on multi-sensor fusion.	Included.	Included.

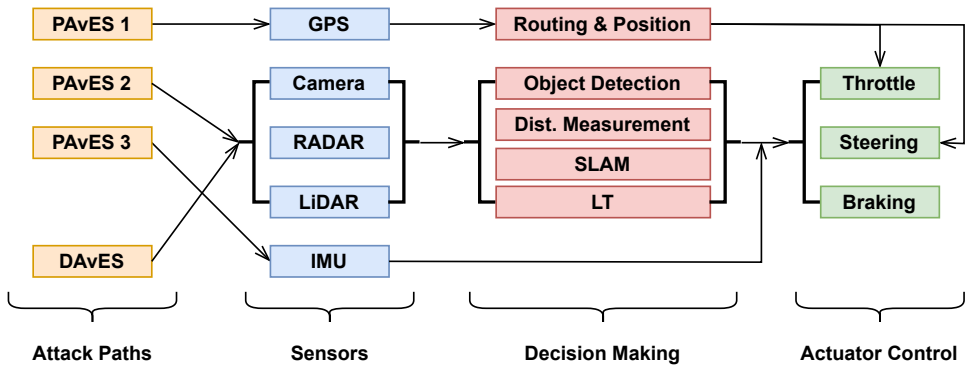


FIGURE 3. Key Components of AV Systems and Their Interaction in Sensor Attack Paths.

the decision making blocks. This can either send incorrect signals to actuators (PAvES 2) or directly influence the control of the actuators (PAvES 2). On the other hand, DAvES involve deploying deceptive objects or adversarial examples to fool sensors, impacting navigational perception data and causing incorrect control signals to the vehicle's steering, throttle, and braking systems.

This section focuses on providing a review of the attacks on navigational perception sensors of AVs. Table 3 gives an overview of different sensor attacks, outlining their goals and navigational perception sensor attack paths. It provides a structured summary of how each attack affects AV systems and the subsequent control signals.

TABLE 2. Major Terms and Definitions Table

Term	Acronym
Adversarial Example	AE
Adversarial Laser Beam	AdvLB
Advanced Driver-Assistance Systems	ADAS
Anomaly Detection based on Point-Level Temporal Consistency	ADoPT
Artificial Intelligence	AI
Automated Lane Centering	ALC
Automatic Dependent Surveillance–Broadcast	ADS-B
Autonomous Driving	DV
Autonomous Vehicle	AV
Average precision	AP
Connected Autonomous Vehicle	CAV
Deep Neural Network	DNN
Electronic Control Units	ECU
Electro-Magnetic-Interference	EMI
Enhanced Realistic Constraints Generation	ERG
Extended Kalman Filter	EKF
False Data Injection	FDI
Fake Shadow Detection	FSD
Feature Interference Reinforcement	FIR
Frequency-Modulated Continuous-Wave	FMCW
Generative Adversarial Network	GAN
Global Navigation Satellite System	GNSS
Global Positioning System	GPS
GPS intrusion detection system	GPS-IDS
Intentional Electro-Magnetic-Interference	IMEI
Inertial Measurement Units	IMU
Light Detection and Ranging	LiDAR
Machine Learning	ML
Micro-Electro-Mechanical Systems	MEMS
Millimeter Wave	mmWave
Multiple Object Tracking	MOT
Multi-Sensor Fusion	MSF
Network Anomaly Detection System	NAD
National Institute of Health	NIH
Packet Capture	PCAP
Pattern-based Attack Classification	PAC
Physical Removal Attacks	PRA
Real-time GPS Signal Generator	RtGSG
Region of Interest	ROI
Robust Physical Perturbations	RP2
Sensor Attack Detection and Classification	SADC
Sequential View Fusion	SVF
Short Lived Adversarial Perturbations	SLAP
Simultaneous Localization and Mapping	SLAM
Time Sensitive Networks	TSN
Time-of-Flight	ToF
True Negative Rate	TNR
True Positive Rate	TPR
Universal Physical Camouflage	UPC
Vehicle-to-Everything	V2X
Virtual Reality	VR
Zero Trust Architecture	ZTA

A. GPS SPOOFING ATTACKS

GPS spoofing attacks pose significant risks to navigation systems, particularly in AVs. These attacks manipulate GPS signals (i.e, transmitting false signals) to lead to inaccurate location, speed, and routing data. Such deceptions can cause AVs to misinterpret their position, potentially triggering false road condition alerts or incorrect hazard detection. Since AVs rely on precise sensor data for vital functions like collision avoidance and lane-keeping, GPS spoofing can result in inappropriate course, speed, or positioning adjustments. This endangers the AV's occupants and jeopardizes the safety of pedestrians and other road users, underscoring a major safety challenge in deploying autonomous navigational perception systems.

Papers on GPS spoofing attacks in autonomous vehicles include [54]–[59]. The first type of attack is Turn-by-Turn Spoofing, where the attacker gradually shifts the GPS position, misleading the vehicle's navigation system into recalculating an incorrect route, potentially leading it to an unintended location [54], [58]. Another attack is the Stop Attack, in which the vehicle is deceived into believing it is in motion while it is actually stationary, causing erroneous decisions at traffic signals or stops [55]. Similarly, the Overshoot Attack manipulates GPS signals to falsely indicate that the vehicle is stationary when it is actually moving, leading to unsafe maneuvers [56]. Additionally, Multi-Sensor Fusion Spoofing targets the vehicle's sensor fusion system, compromising its ability to correctly interpret localization data by altering GPS inputs while keeping other sensors unaltered [57]. Lastly, Machine Learning-Based Spoofing employs adversarial techniques to modify GPS signals in a manner that bypasses traditional anomaly detection algorithms [59].

The research in [60] focuses on GPS spoofing attacks on DJI Phantom 4 Pro drones [61]. The spoofing involves transmitting false GPS signals to the drone, misleading it from its intended path and potentially gaining full control over its navigation. Although the DJI Phantom 4 Pro relies on an inertial measurement unit (IMU) for navigation, it remains susceptible to GPS spoofing attacks. The study underscores the vulnerability stemming from the unencrypted nature of civilian GPS signals, which makes spoofing relatively easy to execute. Using a LabSat3 GPS simulator [62], the researchers designed an experimental setup to demonstrate successful spoofing attacks. The potential outcomes of such attacks include forcing drones into emergency landings, redirecting them to restricted areas, or sending them to unintended destinations. The findings stress the critical need for robust security measures to mitigate these vulnerabilities and protect against the risks posed by GPS spoofing.

Techniques for controlling non-cooperative UAVs through GPS spoofing have been exploited in [13]. The study demonstrates how an attacker can use a low-cost GPS signal simulator based on GNU Radio and an SDR platform to create a false GNSS signal, which helps mislead the drone's position. This attack aims to enforce no-fly zones by tricking UAVs into believing they are outside these zones, causing them to

TABLE 3. Overview of different navigational perception sensor attacks with their methodology and attack paths.

Attack Category	Attack Methodology	Attack Path	Paper
GPS Spoofing	Sends fake GPS signals to deceive GPS receivers into reporting an incorrect location.	PAvES 1	[9]–[11]
	Sends fake GNSS signals to deceive GPS receivers into reporting an incorrect location.	PAvES 1	[12], [13]
LiDAR Spoofing	Emits a laser pulses to manipulate the data received the data received by the LiDAR.	PAvES 2	[14]–[18]
	Manipulates point cloud to alter perceived shapes to deceive the LiDAR system.	DAvES	[11], [19]
	Places an object in the LiDAR field of view to deceive the LiDAR perception system.	DAvES	[20]
Camera Spoofing	Applies sticker pasting to an object to deceive camera sensor and alter input of AV system.	DAvES	[21]–[31]
	Light projection which alters vision system input.	PAvES 2	[32]–[39]
IMU Spoofing	Manipulating signals of IMU to corrupt the motion detection signals in AV and UAV.	PAvES 3	[40]–[44]
Microphone Spoofing	Transmitting inaudible commands to corrupt AV voice recognition systems.	PAvES 1	[45]
	Injecting unauthorized audio signals to corrupt AV navigation systems.	PAvES 1	[46]
Ultrasonic Sensor Spoofing	Manipulating ultrasonic sensor data to corrupt distance measurement in AV systems.	PAvES 2	[15], [47]
mmWave Radar Spoofing	Deceives radar systems that operate in the millimeter-wave frequency band.	PAvES 2	[48], [49]
Multi-Sensor Fusion Spoofing	Manipulating data from multiple sensors in a system designed to aggregate and analyze information from diverse sensor sources.	PAvES 2	[11], [50]–[53]

divert away. The spoofing characteristics include the ability to dynamically change the location information of the spoofing signal relative to the drone's position, effectively steering it away from restricted areas without physical interception.

The authors in [12] presented a method of GNSS spoofing aimed at redirecting non-cooperative drones. The attacker uses a GPS simulator to generate false GPS signals, misleading the drone's navigation system. The attacker seeks to direct the drones to a specified location for remote capture. The spoofing is characterized by its low cost and effectiveness, with an emphasis on being non-destructive to the UAV.

Findings in [9] present GPS spoofing to control unmanned aerial vehicles (UAVs) without physical interception. In the study, the attacker uses a Real-time GPS Signal Generator (RtGSG) for over-the-air spoofing experiments, which deceive the UAV's navigation system, leading to altered trajectories and control of the UAV's movement. The primary objective of these attacks is to gain complete control over the UAV's maneuvers, including its velocity and direction. The spoofing characteristics highlighted include the necessity for real-time manipulation of the GPS signals and the challenges in precisely controlling the UAV's response to these signals, such as maintaining a specific velocity or executing controlled turns.

An investigation into GPS spoofing targeting road navigation systems is documented in [10]. The attacker uses a portable GPS spoofer to generate false GPS signals to reroute victims to unintended destinations without detection. The attack carefully shifts GPS locations to keep navigation instructions consistent with real roads and avoid suspicion. The attack involves creating false GPS signals, shifting the GPS location slightly so that the fake route matches the actual road layout. This approach deceives the navigational perception system without alerting the user, leading them to the wrong destination.

B. LIDAR SPOOFING ATTACKS

LiDAR spoofing attacks pose a significant threat to the perception systems in AVs and robotic systems that rely on LiDAR for environmental mapping and obstacle detection. Since LiDAR systems play a critical role in the safety mechanisms of AVs, their resilience against such attacks is essential for ensuring secure and reliable vehicle operations. Manipulating LiDAR data can severely disrupt path planning and vehicle control systems, leading to safety risk to passengers, pedestrians, and surrounding traffic. In this survey, we classify LiDAR spoofing attacks into three main categories: laser projection, shape manipulation, and object placement.

1) Laser Projection (PAvES 2)

This LiDAR spoofing attack employs direct laser injections to deceive the LiDAR system into perceiving objects at

incorrect distances or positions. Such precision manipulation of LiDAR data can mislead an AV's path planning, potentially causing unsafe maneuvers or failure to recognize actual hazards.

The study in [18] investigates the susceptibility of AV systems to LiDAR spoofing attacks. By injecting fake data points, the attacker alters the vehicle's perception of its surroundings, tricking it into perceiving non-existent obstacles and making erroneous driving decisions. This sophisticated attack leverages adversarial machine learning techniques to optimize the spoofed data for maximum impact, highlighting the critical need for robust security measures in AV perception systems.

Further exploration in [17] reveals the vulnerability of LiDAR systems to physical laser attacks, where a laser transceiver injects numerous spoofing points into the LiDAR system. This deception aims to disrupt the 3D object detection system, either by hiding real objects or creating fictitious ones. Extensive experimental evidence highlights the necessity for robust countermeasures in LiDAR-based AV navigational perception systems.

The research presented in [16] addresses black-box adversarial sensor attacks on LiDAR-based perception systems. Attackers manipulate LiDAR data to fabricate false objects or conceal real ones, leading to potentially hazardous driving decisions. This study uncovers a vulnerability related to the occlusion patterns in point clouds and demonstrates high success rates across different model designs, emphasizing the urgent need for resilient countermeasures.

Additionally, [14] examines laser-based spoofing attacks on LiDAR systems, demonstrating how laser pulses can remove point cloud data of actual obstacles, thus preventing AVs from recognizing these obstacles. By exploiting LiDAR filtering processes, such attacks pose significant safety risks, underscoring the imperative for enhanced defense strategies against these sophisticated threats.

Furthermore, Yi et. al [63] proposed a feature-based spoofing attack strategy that causes localization drift in LiDAR-based localization algorithms. The spoofing attack involves an attacker's transceiver intercepting the original signal and sending a counterfeit laser pulse to deceive the LiDAR system. During this process, two sequential pulse generators transform the input into a fake signal, with variations in the number of periods, delay times, pulse quantities, and copies. This strategy can deceive LiDAR without interfering with the processing or transmission of the sensor measurements. Specifically, the feature-based attacker extracts features from the point cloud generated by perceiving the surrounding environment, similar to how a LiDAR SLAM system operates. The attacker then calculates optimized perturbations on these extracted feature points and directs the LiDAR spoofing devices to alter the original feature point positions, causing them to shift and become fake outliers. These fake outliers are then injected into the LiDAR sensor as spoofed reflections. As a result, the victimized LiDAR SLAM system is misled by the corrupted input point cloud and produces incorrect pose estimations. Consequently, the trajectory drifts in the presence of the feature-based attack.

Moreover, Li et. al [64] proposed Fool LiDAR perception with Adversarial Trajectory (FLAT), which adds perturbations to vehicle trajectories. This can result in a significant drop in the precision of the object detector, potentially reducing it to nearly zero.

2) Shape Manipulation (DAvES)

In this LiDAR spoofing attack, the adversary alters the perceived shapes of objects, severely distorting the vehicle's environmental understanding. The study in [11] examines the vulnerability of AV perception systems to physical-world attacks that targets both camera and LiDAR sensors. By creating a stealthy 3D-printed object that deceives both sensor types, the vehicle fails to detect the object, significantly increasing collision risk. This research highlights the susceptibility of multi-sensor fusion systems to coordinated physical attacks, emphasizing the need for robust defense strategies in AV technology.

The research in [18] explores the vulnerabilities of LiDAR-based detection systems in AVs to adversarial attacks. It introduces LiDAR-Adv, an optimization-based approach to generate adversarial objects that evade detection under various conditions. The attack involves simulating a differentiable LiDAR renderer and formulating 3D feature aggregation with a differentiable proxy function. The resulting image is then 3D printed and placed in the environment to replace the original objects. The objective is to either conceal an existing object or alter the classification of a detected object by manipulating its shape, thereby misleading the detection system.

3) Object Placement (DAvES)

This LiDAR spoofing attack involves generating fake LiDAR reflections to create the illusion of non-existent objects in the sensor's field of view. Such attacks could cause AVs to take evasive actions to avoid these phantom obstacles, potentially leading to real accidents or traffic disruptions.

The research in [20] explores the feasibility of using random objects with reflective surfaces to deceive LiDAR perception in AVs. The attack mechanism revolves around placing these objects at strategic locations that are determined to be most effective in distorting the LiDAR's navigational perception. It is shown that this could lead to the LiDAR system failing to accurately detect real objects, such as other vehicles, posing a significant safety risk. A key characteristic of this attack is its reliance on everyday reflective objects, making it a practical and potentially easy-to-execute method for disrupting LiDAR systems in AVs.

4) Other Techniques

Unlike the aforementioned LiDAR spoofing attacks, there are other attacks that target the LiDAR detection runtime, such as the one proposed in [65]. The authors propose a novel attack technique aimed at increasing the runtime latency

of LiDAR-based object detectors. This attack utilizes point injection and point perturbation methodologies, combined with a unique loss function that quantifies the impact of adversarial corruption on the input relative to the network's execution time.

C. CAMERA SPOOFING ATTACKS

Camera spoofing attacks pose a significant threat to AVs and robotic systems using cameras for navigation and object recognition. Like LiDAR spoofing, these attacks can misinterpret visual data, leading to dangerous outcomes. The two primary types of camera spoofing attacks are sticker pasting and light projection.

1) Sticker Pasting (DAvES)

In camera spoofing sticker pasting attacks, physical stickers are placed on traffic signs, vehicle exteriors, or other parts of the environment that the AV's cameras can detect. These stickers can alter the appearance of signs, causing misclassification, or can create fictitious signals that the camera system interprets as real. This could result in the AV ignoring traffic laws or reacting to non-existent road rules, thus endangering its occupants and other road users.

A study using the CARLA simulator has found that stop signs with stickers that are visually camouflaged can substantially hinder or prevent proper traffic sign detection [31]. These minimal yet adversarial modifications illustrate how simply an attacker can leverage physical-world attacks to exploit camera-based perception. Similarly, the work in [22] introduces a novel attack method against AVs' Automated Lane Centering (ALC) systems. It fools deep neural network-based ALC systems with dirty road patches as adversarial perturbations. Such inconspicuous patches are less noticeable to ALC systems and drivers since they look like natural road dirt and are crafted to be stealthy. The attack is meant to induce large lateral deviations rapidly, thereby steering the vehicle off its lane within less than the average driver reaction times. This approach demonstrates a critical vulnerability in ALC systems and a need for more security in autonomous driving technology.

The research in [23] introduces an innovative method for creating adversarial attacks against deep learning models, particularly focusing on autonomous driving systems. The attack mechanism involves generating adversarial camouflage that can deceive both machine models and human perception. This is achieved by suppressing attention in deep learning models and evading human visual attention. The goal is to create camouflages that blend seamlessly into the physical environment, thereby misleading automated systems without arousing human suspicion.

The research in [24] presents a novel adversarial attack against the Multiple Object Tracking (MOT) system in autonomous driving. This attack, termed 'tracker hijacking,' involves creating adversarial examples that mislead the object tracking process. The goal is to manipulate the tracking error reduction process in MOT, allowing attackers to alter objects' trajectories in an AV's perception system. This can lead to safety hazards as the car might fail to properly track and react to other vehicles or obstacles on the road.

The research in [30] introduces a novel method to create adversarial attacks against autonomous driving systems. The attack employs a Generative Adversarial Network (GAN) framework to produce adversarial examples that can continuously mislead an autonomous driving model's steering mechanism. These examples, designed to resemble roadside signs, are visually indistinguishable from real-world objects and are designed to be robust against physical-world conditions. The primary goal of this attack is to generate a single adversarial example that can effectively and consistently mislead the steering model throughout the driving process.

The research in [29] explores the development of physical adversarial attacks on object detection systems used in AVs. The attack involves applying visually subtle perturbations to physical objects, such as stop signs, to render them undetectable or misclassified by object detection algorithms. The primary objective is to highlight the vulnerability of these systems to real-world adversarial manipulations, which could mislead autonomous driving systems and create potentially hazardous situations.

The research in [28] introduces the Robust Physical Perturbations (RP2) algorithm. This algorithm generates physical adversarial perturbations designed to fool deep neural network classifiers under real-world conditions. The paper focuses on misclassifying traffic signs by applying visually subtle modifications, such as stickers, which result in incorrect classifications by the targeted system. The goal of these attacks is to demonstrate that even minor physical changes to objects like road signs can mislead advanced visual classification systems, posing significant implications for AV safety.

The research in [25] focuses on developing robust adversarial examples (AEs) to deceive real-world object detectors in autonomous driving. It introduces techniques like feature interference reinforcement (FIR) and enhanced realistic constraints generation (ERG) to improve the robustness of AEs. The study addresses both Hiding Attacks, which render objects undetectable, and Appearing Attacks, which cause objects to be misidentified. The findings demonstrate that these AEs can effectively deceive state-of-the-art object detectors, including YOLO V3 and Faster R-CNN, with a high success rate, even under varying distances, angles, and lighting conditions.

The research in [26] investigates the vulnerability of AVs' lane detection systems to crafted perturbations. The study reveals that even small, manually added road markings can deceive these systems. The researchers propose a two-stage approach for generating such markings, which includes determining optimal perturbations on a camera image and then mapping them to physical road markings. Their experiments, conducted on a Tesla Model S, show that these subtle perturbations can successfully mislead the vehicle's lane detection module, causing it to drive in the wrong direction.

The research in [27] introduces a method for creating adversarial attacks that are effective in the physical world against object detection systems. It proposes the Universal Physical Camouflage (UPC) technique, which crafts patterns that can universally deceive object detectors into either failing to recognize objects or misclassifying them. The attack is notable for its ability to work on various object types and in different physical conditions, making it a versatile and potent method for undermining object detection systems.

The research in [66] considers placing crafted stickers on cameras to cause object misidentification. Specifically, the authors introduced the translucent patch attack. This attack involves placing a translucent patch, optimized through gradient-based methods, onto the camera's lens to obscure instances of a specific target object class, such as stop signs, while still allowing the detection of other untargeted objects. Unlike traditional adversarial attacks that require direct access to the object being manipulated, this attack affects the camera sensor itself, enabling a more practical and stealthy approach. The experiments show a significant reduction in detection accuracy, with the attack decreasing the average precision of stop signs by over 42% in both digital and physical settings. The patch remains difficult for human observers to notice, making it an effective method to deceive state-of-the-art object detectors, such as YOLO and Faster R-CNN, without compromising the detection of other road objects like cars or pedestrians

2) Light Projection Attack (PAvES 2)

In a camera spoofing light projection attack, the attacker uses a projected light to create illusions or obscure the AV's camera vision. By projecting images onto surfaces or directly into the camera's lens, attackers can trick the vehicle's vision system into seeing objects that aren't there or into missing real objects that are present. Such attacks could cause the AV to navigate based on false information, such as nonexistent lane markings or phantom obstacles, leading to improper navigational perceptions.

The research in [38] introduces a new type of physical-world adversarial attack using a laser beam. This attack, named AdvLB (Adversarial Laser Beam), uses a laser to project adversarial perturbations onto objects, causing deep neural networks (DNNs) to misclassify them. The goal is to exploit the speed and flexibility of lasers to perform fast and effective attacks. The paper demonstrates that this method, with high success rates, can deceive DNNs in both digital and physical settings.

The research in [39] aims to use a new form of physical-world adversarial attack: focused laser beams to compromise camera sensor reliability. By drowning the camera with bright or modulated light, this attack actually "blinds" the sensor, overloading or distorting its feed and causing missing or false detections. The goal is to exploit optical vulnerabilities in vision-based perception algorithms, thereby obscuring crucial objects and degrading real-time vehicle guidance. The paper demonstrates that such laser-based interference poses a substantial risk, especially in scenarios where accurate camera data is critical for collision avoidance and lane-keeping.

The research in [37] explores methods to manipulate unmanned aerial vehicles (UAVs) by spoofing optical flow sensors. The primary attack mechanism involves altering the perceived environment of these sensors, which can deceive the UAV's navigation system, leading to incorrect movements. The goal is to gain control over the UAV's lateral movement, potentially causing it to collide with obstacles. This approach underscores the need for more secure navigational perceptions systems in UAVs to prevent such manipulation.

The research in [36] presents an attack method targeting AVs' and drones' depth estimation systems. The attack, named DoubleStar, involves using two light sources to create false depth perceptions by exploiting stereo camera vulnerabilities. This method causes the systems to misinterpret the depth of objects, potentially leading to collision avoidance failures. The unique aspect of this attack is its long-range capability and the use of common light sources, making it practical and challenging to detect.

The research in [35] explores a novel attack method against driver-assistance systems. This method, refered to as "split-second phantom attacks". This approach uses brief visual projections or digital displays to trick these systems. The attack aims to create illusions of obstacles or road signs, causing the vehicle to react unexpectedly, like sudden braking or issuing false notifications. These projections are designed to be detected by the system but remain invisible to human drivers.

The research in [34] explores the feasibility of manipulating traffic light recognition systems using laser interference. The researchers use a laser to generate color stripes on traffic lights in camera images with rolling shutters. This method can trick a traffic light recognition system into misinterpreting the light's color, for instance, seeing a red light as green or vice versa.

The research in [21] explores remote attacks on camera-based image classification systems. It utilizes adversarial patterns projected into cameras, exploiting optical effects like lens flare. In this attack, the attacker projects adversarial patterns into camera systems by exploiting optical effects like lens flare and auto-exposure control. The attacker's main goals include creating spurious objects (objects that don't exist in reality) in the camera's perception and altering the appearance of existing objects to be recognized as something else. These attacks are stealthy, require no physical access, and can deceive systems by altering object appearances remotely.

3) Other Camera Spoofing Techniques

Some adversarial attacks target detection models by introducing adversarial examples in which input images input images are modified with minute, nearly undetectable noise. For example, the research in [67] introduces an attack model that uses adaptive Adversarial Billboards to override the control of autonomous driving systems that rely on DNNs. This

attack is physically realizable and involves placing a digital billboard by the roadside that displays dynamically changing adversarial images or videos. The billboard is equipped with a camera to track the approaching vehicle's position (pose) and adjust the displayed content in real-time, adapting to the vehicle's motion and environmental conditions such as lighting and weather. By manipulating the vehicle's perception system, the billboard images cause the DNN to output incorrect steering commands, guiding the vehicle to follow a trajectory specified by the attacker instead of its original path. The attack was tested in the CARLA simulator, a high-fidelity platform for autonomous driving, and demonstrated the ability to force autonomous vehicles to deviate from their intended route, such as making a car turn left when it should have gone straight or steering it toward an adversarial path. The results show that this method is highly effective under various scenarios, including different weather conditions and traffic, making it a robust threat to the safety of autonomous vehicles.

The research in [68] presented backdoor attacks that target lane detection systems in autonomous vehicles by poisoning the training data with adversarial triggers. These triggers, designed as traffic cones with specific shapes and positions, are introduced into the training set through two methods: poison-annotation and clean-annotation. In the poison-annotation attack, lane boundaries are deliberately misannotated to mislead the vehicle off course. The clean-annotation attack exploits image scaling vulnerabilities to conceal the adversarial trigger while maintaining visually correct annotations. These attacks are model-agnostic, meaning they can compromise various lane detection algorithms. Once the backdoor is activated, the vehicle misinterprets the lane markings, potentially leading to dangerous driving behaviors, such as veering into the wrong lane. The experiments demonstrated the effectiveness of these attacks in both simulated and real-world settings, posing significant safety risks to autonomous driving systems.

D. IMU SPOOFING ATTACKS (PAvES 3)

IMU spoofing attacks represent a significant and insidious risk to the inertial navigation systems of AVs and robotic platforms. The IMU is a critical sensor that provides essential data on a vehicle's acceleration and rotation, influencing its understanding of movement and orientation in space. Spoofing an IMU can cause errors in vehicle navigation and stability control systems. IMU spoofing is particularly hazardous because it targets fundamental aspects of the vehicle's operation, which are generally assumed to be reliable.

The research in [40], [44] investigates the manipulation of embedded Micro-Electro-Mechanical Systems (MEMS) inertial sensors in various devices through non-invasive acoustic attacks. These attacks use sound waves at specific frequencies to induce resonant vibrations in the sensor, resulting in false readings. The goal is to gain implicit control over the system's actions without physical contact, a method that can impact a wide range of devices, from virtual reality (VR) systems to navigation and self-balancing transporters.

The research in [43] examines how acoustic signals can manipulate inertial sensors in AV object-detection systems, leading to misclassification. The Poltergeist attack targets image stabilizers, which compensate for camera jitters. By emitting designed acoustic signals, an adversary can control the output of these sensors, causing unnecessary motion compensation and resulting in blurred images. These blurred images can mislead object detection algorithms, affecting critical safety decisions.

The research in [42] investigates the vulnerability of drones to acoustic attacks targeting their gyroscopic sensors. The study demonstrates how specific sound frequencies can induce resonant vibrations in the sensors, causing false readings. This manipulation can cause a loss of control over the drone's movement, potentially resulting in crashes.

The research in [41] examines vulnerabilities in sensor fusion systems that use Kalman Filter-based algorithms, specifically focusing on how these systems can be compromised via signal injection attacks. The paper presents two types of attacks: stationary and non-stationary. The stationary attack manipulates the accelerometer and magnetometer, misleading the system into a false stationary state. In contrast, the non-stationary attack involves manipulating the gyroscope, often coupled with disturbances to the accelerometer and magnetometer, creating a misleading state of motion. As the Kalman Filter algorithm is used to fuse data from multiple sensors (accelerometer, gyroscope, magnetometer), it employs a mathematical approach to estimate the true state of a system by considering the probability of the errors in sensor readings for estimating the hidden states from the erroneous state. These attacks exploit the Kalman Filter's reliance on sensor inputs, and when these inputs are corrupted, they can lead to incorrect estimations of the system's state. The researchers demonstrate how false sensor signals can manipulate the fusion output on inclination measurements in altitude-heading reference systems (AHRS), leading to erroneous inclination readings. This highlights the critical need for robust security measures in systems utilizing Kalman Filters for sensor fusion.

E. ACOUSTIC SENSOR SPOOFING ATTACKS (PAvES 1)

Microphone spoofing attacks present a sophisticated threat to systems that rely on acoustic sensors for voice command recognition, environmental awareness, and communication. These attacks target devices' microphones, such as AVs, smart home devices, and voice-controlled systems, and can lead to serious security breaches and malfunctions. We categorize microphone spoofing attacks into inaudible voice attacks and audio injection attacks.

1) Inaudible Voice

The research in [45] explores a method called "DolphinAttack" to control voice assistants like Siri and Google Now using ultrasonic frequencies. This method involves modulating voice commands onto ultrasonic carriers, making them

inaudible to humans but still perceivable by voice assistants. The goal is to silently control these devices to perform unauthorized actions, such as making phone calls or opening websites, without the user's knowledge or even sending wrong destination to the GPS voice prompt.

2) Audio Injection

The research in [46] investigates how to compromise systems using Kalman filters, commonly found in AVs, with signal injection attacks. It demonstrates that injecting malicious signals into sensory inputs makes it possible to manipulate the filter's output, leading to incorrect system responses. The goal is to show that even systems relying on advanced filters like Kalman are vulnerable to carefully crafted interference.

F. ULTRASONIC SENSOR SPOOFING ATTACKS (PAvES 2)

Ultrasonic sensor spoofing attacks target the ultrasonic range sensors widely used in various navigational perception of AVs' robotic systems. They are especially used for and proximity sensing. These sensors use high-frequency sound waves to detect objects and measure distances, which is crucial in navigation, collision avoidance, and parking assistance systems.

The research in [47] focuses on the vulnerabilities of ultrasonic sensors in vehicles to spoofing and jamming attacks. It demonstrates that these sensors, crucial for parking assistance and collision avoidance, can be deceived into detecting nonexistent obstacles or failing to detect real ones. The study uses various techniques, such as random spoofing, adaptive spoofing, and jamming to test these vulnerabilities.

The research in [15] examines the vulnerability of AV sensors to contactless attacks. In this attack, attackers exploit ultrasonic sensors, millimeter-wave radars, and cameras, which are critical for the vehicle's navigation and safety systems. The study demonstrates that these sensors can be jammed or spoofed using off-the-shelf hardware, leading to dangerous situations like crashes or impaired vehicle safety.

G. mmWAVE RADAR SPOOFING ATTACKS (PAvES 2)

Millimeter wave (mmWave) radar spoofing attacks are an emerging threat targeting the millimeter-wave radar systems used in AVs and advanced driver-assistance systems (ADAS) (usually in Level 3 and above). mmWave radars are critical for detecting objects, measuring distances, and determining the relative speed of surrounding objects, especially in adverse weather conditions where other sensors like cameras and LiDAR may be less effective.

The research in [69] investigates a method for deceiving mmWave Frequency-Modulated Continuous-Wave (FMCW) radars used in AVs. The attack employs a low-cost setup, which includes a commercial 24 GHz mmWave FMCW radar module, an analog device, and Arduino Due. This setup is used to replicate and replay radar signals, causing delay offset and generating fake distance readings. Such manipulation can mislead the vehicle's perception system, potentially leading to unsafe driving decisions.

The research in [49] focuses on the security of mmWave radar systems in AVs. It presents practical attack methods in which attackers spoof the radar system by generating false signals, leading to the detection of non-existent obstacles or the misplacement of real ones. The goal is to manipulate the AV's decision-making process, posing risks like incorrect obstacle avoidance or navigation decisions. The study demonstrates these attacks in real-world scenarios and proposes defense mechanisms, including challenge-response authentication and RF fingerprinting, to mitigate such threats.

The research in [48] proposes using 3D-printed objects to attack deep learning-based radar perception models. Specifically, the 3D-printed objects are reflective and manipulate signal reflectivity to deceive DNN-based detection models, rendering them unable to detect targets.

The research in [70] presents a spoofing attack designed to deceive the radar system into miscalculating the actual distance to objects or detecting non-existent obstacles. This is achieved by overwriting the actual reflection with a spoofing signal that is coherent with the victim's transmitted signal, sharing the same chirp rate but at a higher power.

The research in [71] considers a radar information inference system using COTS radar to detect the mode and estimate the parameters of a target radar, which can be employed to facilitate mmWave radar spoofing.

H. MULTI-SENSOR SPOOFING ATTACKS (PAvES 2)

Multi-sensor fusion spoofing attacks pose a complex challenge to the integrated systems within AVs. In such cases, multiple sensors are combined to create a comprehensive understanding of the vehicle's environment. This sensor fusion is fundamental for AVs to make accurate and safe navigational decisions. Spoofing attacks that target multi-sensor fusion systems aim to manipulate the data from several types of sensors simultaneously, exploiting how these systems integrate information to create a coherent world model. Such attacks aim to create inconsistencies or false alignments in the data being processed, leading to potential misjudgments by the AV.

The research in [52] explores the vulnerability of sensor fusion systems in AVs to single-modal attacks. In this attack, the attacker targets the AV camera-LiDAR fusion models. These models combine camera data and LiDAR sensors inputs for tasks like 3D object detection. The attack is single-modal, targeting only the camera modality. The attacker uses a two-stage approach. First, attackers identify sensitive image areas vulnerable to adversarial attacks. Then, they apply customized strategies to create deployable patches that manipulate these sensitive areas. The attacks lead to compromised object detection capabilities; causing it to miss or incorrectly identify objects. This, in turn, impacts the AV's decision-making and control systems. As a result, the attacker can significantly reduce the model's object detection performance, leading to scenarios where the AV might fail to

recognize important objects like other vehicles or pedestrians, potentially causing hazardous situations.

The research in [11] investigates the vulnerabilities of AV perception systems to physical-world attacks. It demonstrates that both camera and LiDAR sensors can be deceived simultaneously, leading to a failure in object detection. The attacks involve adversarial 3D objects that are specifically designed to evade detection by manipulating their shape to distort the inputs of both camera and LiDAR sensors. The attackers manipulate the shape of 3D objects to generate inconsistencies in sensor readings. By altering the shape of these 3D objects, the attackers create inconsistencies in sensor readings, which disrupt the AV's perception system, causing incorrect or failed object detection. As a result, the AV's ability to perceive and respond to real-world obstacles is compromised, significantly increasing the risk of collisions. The attack uses a unique approach to create a 3D-printed object that is stealthy enough to deceive both sensor types, preventing the AV from detecting it.

The research in [51] investigates the security of multi-sensor fusion (MSF) systems in AVs against GPS spoofing attacks. The study discusses an MSF approach that integrates inputs from GPS, IMU, and LiDAR. The focus is on GPS spoofing attacks, where false GPS signals are used to mislead the localization system of AVs. Spoofing attacks primarily target the GPS sensor. Such attacks can lead to incorrect localization data, potentially causing the vehicle to deviate off the road or into wrong-way driving, thus impacting decision-making and control systems. The attacks exploit vulnerabilities in the sensor fusion process, especially when the MSF system is less confident, which can result in a take-over effect where the spoofed GPS data becomes the dominant input in the fusion process. This undermines the MSF principle of leveraging multiple sensors for robustness and accuracy. The research demonstrates that while MSF generally enhances security against GPS spoofing, vulnerabilities still exist due to dynamic and non-deterministic factors like sensor noises and algorithm inaccuracies. The paper introduces "FusionRipper," a novel attack exploiting these vulnerabilities, achieving high success rates in causing off-road and wrong-way deviations.

The research in [50] examines the vulnerability of AVs to LiDAR spoofing attacks, focusing on the robustness of camera-LiDAR sensor fusion against these attacks. It introduces a novel attack method called the "frustum attack," which is effective against various perception algorithms used in AVs. This attack works by exploiting the inconsistency between the LiDAR and camera data, placing spoofed points in a way that preserves semantic consistency between these two data sources. The paper demonstrates that while specific perception algorithms are robust against naïve LiDAR spoofing attacks, they are significantly vulnerable to the more sophisticated frustum attack. Furthermore, the study reveals that defenses against LiDAR spoofing are ineffective in detecting and mitigating frustum attacks.

The research in [11] investigated adversarial attacks in the context of multi-sensor fusion, aiming to create adversarial examples that are invisible to both camera and LiDAR sensors. Specifically, the authors designed a physical-world adversarial attack targeting AVs by creating 3D objects that could mislead the vehicle's detection systems. The goal was to make the AV fail to detect the adversarial object, potentially leading to a crash. To achieve this, the authors employed an optimization-based approach that addressed several design challenges. These challenges included the non-differentiability of the target camera and LiDAR sensing systems, as well as the non-differentiability of cell-level aggregated features used by LiDAR. The optimization approach they proposed was carefully crafted to overcome these complexities, ultimately producing an adversarial 3D object capable of evading detection by both camera and LiDAR sensors.

The research in [72] demonstrates the feasibility of attacking MSF-based perception by targeting only camera-based perception. Specifically, the authors analyzed the sensitivity and susceptibility distribution of image areas while the network made predictions using gradient information. Based on this analysis, they categorized the fusion networks as either object-sensitive or globally sensitive, embedding the adversarial patch in the surrounding environment or attaching it to the desired object, respectively. Additionally, the authors showed that input transformation-based approaches were ineffective against adversarial patch attacks on MSF networks. They also explored how the network architecture impacts adversarial robustness.

Finally, the work in [73] studies fuzzy-model-based lateral control for networked autonomous vehicle systems that are subject to hybrid cyber-attacks. This research develops resilient adaptive fuzzy control strategies to defend against multiple types of cyber-attacks that could influence sensor data and control signals in networked autonomous vehicles. It provides stability guarantees and performance analyses despite these hybrid attacks, highlighting the importance of robust control design in the presence of cyber threats. Such control frameworks form a critical part in maintaining reliable vehicle behavior under adversarial conditions, complementing efforts that address sensor spoofing and perception attacks.

III. COUNTERMEASURES

In this section, we provide a summary of sensor attack countermeasures and their limitations, as outlined in Table 4. Then, we present a practical analysis to evaluate these countermeasures based on three key metrics: cost, deployment overhead, and scalability. The goal of this analysis is to assess the real-world applicability of each countermeasure by examining not only its theoretical effectiveness but also its feasibility for integration into existing AV/RV platforms. These metrics are crucial for understanding the trade-offs between protection level and implementation burden, as detailed in Tables 4 and 5. Our evaluation of the countermeasure cost includes a review of relevant literature, focusing on hardware

requirements (ranging from low-cost to high-cost solutions) and software demands, such as development complexity and required modifications to existing code bases.

A. COUNTERMEASURES AGAINST GPS SPOOFING

Several techniques have been proposed as countermeasures against GPS spoofing attacks, which can be categorized into two main approaches: 1) modification-based and 2) modification-free, as described in [10].

A-1: Modification-Based Approaches

Modification-based approaches can be classified into: encryption and signal handling; ground infrastructure; and GPS receiver hardware modifications.

- *Encryption and signal handling*: This strategy involves enhancing civilian GPS signals with P(Y) code encryption and integrating signal authentication into the forthcoming Global Navigation Satellite System (GNSS) generation. Despite its potential, this approach is impeded by the extensive number of existing civilian GPS devices and the logistical challenges associated with rapidly implementing these upgrades. Several research works proposed encryption and signal handling schemes for GPS spoofing mitigation. For example, the research in [78] focuses on enhancing GPS signal security through encryption and authentication. This method uses "hidden markers" in navigation signals and cryptographic keys broadcasted with a delay. This technique separates the ability to authenticate a signal from the capability to generate one, making spoofing more difficult. However, the approach has limitations, particularly in defending against selective-delay attacks using directional antennas. It utilizes secure bit-sequence generators and affordable converters. It requires significant signal structure and receiver changes, implying a system-wide update. It can be easily integrated into multiple systems, but maintaining cryptography may pose a challenge. Additionally, the research in [79] proposes a technique combining cryptographic authentication of the GPS navigation message with signal timing authentication. It advocates using digital signatures to ensure the integrity and origin of the GPS signals, thereby increasing the difficulty of executing a successful spoofing attack. However, the method is not entirely immune to sophisticated attacks, and involves significant complexity in implementation, requiring modifications to the existing GPS infrastructure.
- *Ground infrastructure enhancement*: Adopting reliable ground infrastructures aids GPS devices in authenticating their location. Methods include trusted verifiers, distance-bounding protocols, multiliterate GPS, multi-receiver crowdsourcing, and checks at the physical layer. However, governmental policy constraints and significant financial investments are major hurdles in deploying these infrastructures. Several research works employed ground infrastructure enhancement as GPS spoofing countermeasures. For instance, the research in [87] focuses on determining the maximum physical distance between a receiver and a transmitter to prevent relay attacks, a form of spoofing. However, the accuracy of timing measurements and environmental factors can limit this method. The research in [80] introduces a system leveraging crowdsourced data from numerous ground sensors to detect and localize GPS spoofing attacks on aerial vehicles. This system does not require changes to GPS infrastructure or airborne GPS receivers. It detects spoofing by analyzing the contents and arrival times of Automatic Dependent Surveillance–Broadcast (ADS-B) and FLARM [88] position advertisements from aircraft affected by spoofing. While the approach allows for rapid spoofing detection and attacker localization, its effectiveness depends on the density and distribution of the sensor network. Also, the research in [89] examines the threat of multi-device spoofing attacks on air traffic control and suggests countermeasures like analyzing physical-layer features of signals. However, these measures require complex detection systems and may not be entirely foolproof. Similarly, the research in [81] investigates radio frequency distance bounding as a method to authenticate the distance of a signal source. This approach mitigates spoofing by verifying the signal's origin but encounters challenges, including signal interference and the need for precise timing mechanisms. Reference [82] proposes verifying the integrity of tracking data as a means to mitigate spoofing risks. While this technique is theoretically effective, its practical implementation is challenging due to the variability of data sources and the need for robust verification algorithms.
- *GPS receiver hardware modifications*: Another approach involves modifying GPS receivers to incorporate features like signal angle-of-arrival for authenticity verification. For example, the research in [90] utilizes directional antennas to mitigate the risk of wormhole attacks in network communications. While effective, this approach is tailored to specific network configurations. Similarly, the research in [83] proposes robust position estimation in wireless sensor networks to enhance location accuracy and resist spoofing. Although theoretically effective, this method may encounter practical implementation challenges. The research [84] explores the use of multi-antenna receivers to detect spoofing, enhancing detection capabilities but increasing receiver complexity. Reference [85] employs high-frequency antenna motion and carrier-phase data for GNSS spoofing detection, offering effective results but involving significant technical complexity. Reference [86] introduces the quickest detection algorithm for GPS spoofing, aiming for rapid detection with minimal false alarms, striking a balance between speed and accuracy. However, the feasibility of these methods is limited by the wide range of mobile devices currently in use, which would require

TABLE 4. Existing countermeasures to AV navigational perception sensor attacks & their limitations.

Attack Type	Countermeasure Techniques	Limitations and Practical Feasibility
GPS Spoofing [10]	**Modification-based approaches:** - Encryption and signal handling. - Ground infrastructure enhancement. - GPS receiver hardware modifications. - GPS receiver software upgrades. **Modification-free approaches:** - External location verification. - Internal sensor fusion. - Computer vision-based location verification. **Dataset:** NYC Taxi and Limousine Commission (TLC) data [74].	**Modification-based approaches:** - Challenges in rapid implementation. - High costs and policy constraints. - Limited feasibility for existing devices. - Susceptibility to advanced spoofing. **Modification-free approaches:** - Prone to IMU sensor errors. - Vulnerability to adversarial manipulations.
LiDAR Spoofing - Laser Projection [16], [50]	FSD. Azimuth-based Detection. CARLO. SVF. **Dataset:** KITTI [75].	FSD: incurs high computational complexity and may cause delays in real-time applications. Azimuth-based Detection: it is not applicable to adversarial spoofing or other ORA-based attacks. CARLO: does not require model retraining. SVF: requires model retraining. Both: Potential issues with future increases in sensor attack capability; small portion of false alarms.
LiDAR Spoofing - Shape Manipulation [11]	**Multi Sensor Fusion :** Combining LiDAR data with camera or radar inputs to improve detection accuracy and filter out manipulated shapes. **Adversarial Object Detection:** Detecting inconsistencies in object shapes using adversarial patterns. **Physical Scene Verification:** Utilizing reference structures or prior scene knowledge to validate detected objects. **Dataset:** KITTI [75].	**Multi Sensor Fusion:** Increases computational load and hardware costs. **Adversarial Object Detection:** It requires robust adversarial training datasets, which may not generalize well. **Physical Scene Verification:** Limited by environmental changes and dynamic scenarios.
LiDAR Spoofing - Object Placement [20]	Adversarial Location Defense. Adversarial Training. Sensor Fusion. **Dataset:** KITTI [75].	Needs balance defense effectiveness, which impacts accuracy. Not effective against some attacks. Cost and complexity.
Camera Spoofing - Sticker Pasting [22]	Machine learning model level defenses (e.g., JPEG compression, bit-depth reduction, adding Gaussian noise, median blurring, and autoencoder reformation). **Dataset:** comma2k19 [76].	While these methods effectively reduce attack success rates under certain parameters, they also lower success rates in benign cases. When maintaining 100% benign-case success rates, attack success rates remain high (99% to 100%). Less effective against physical-world attacks involving human perceptible perturbations. There is a need for novel adaptations of advanced defenses or the development of new defenses tailored specifically to LD.
Camera Spoofing - Light Projection [33]	SentiNet. Adaptive defender using adversarial learning. **Dataset:** GTSRB [77].	SentiNet: SLAP can bypass SentiNet over 95% of the time. Adaptive defender using adversarial learning: Can prevent most attacks but reduces accuracy in non-adversarial conditions.

substantial hardware modifications.

- *GPS receiver software upgrades*: This involves developing algorithms for consistency checks to detect anomalies in GPS signals and enhancing GPS receivers to connect with more satellites or to synchronize with other receivers like the work presented in [91], [92]. Nonetheless, these techniques might be susceptible to advanced spoofing methods.

Table 5 provides a comparative summary of prominent GPS spoofing countermeasures using a qualitative assessment across three key dimensions: cost, deployment overhead, and scalability. Due to the diverse nature of existing literature and the lack of standardized performance metrics such as detection accuracy or latency, a direct quantitative comparison was not feasible. Instead, we adopted a pseudo-quantitative classification approach to offer a structured, practical evaluation. Here, cost is categorized into four aspects: low-cost hardware (e.g., off-the-shelf components for prototyping), high-cost hardware (e.g., custom or high-performance devices), software development (bespoke system implementation), and software upgrade/modification (extending existing software). The deployment overhead reflects the ease with which a countermeasure can be integrated into current AV systems, considering factors like compatibility and physical space. Here, scalability refers to the

TABLE 5. Summary of GPS spoofing countermeasures in terms of cost, deployment overhead, and scalability

Paper	Cost				Deployment Overhead			Scalability
	Low-Cost Hardware	High-Cost Hardware	Software Development	Software Upgrade/Modification	Easy Integration to the Existing System Required	Dynamic Recalibration	Software Change Required	
[60]	X	-	-	X	X	-	-	X
[78]	-	X	-	-	X	-	-	X
[79]	-	-	-	X	-	-	-	-
[80]	X	-	-	-	X	-	-	X
[81]	-	X	-	-	-	X	-	-
[82]	X	-	-	-	-	-	X	-
[83]	X	-	-	X	-	-	-	-
[84]	X	-	-	-	-	-	-	X
[85]	-	X	X	-	-	-	X	-
[86]	X	-	-	-	X	-	-	X

ability of a countermeasure to maintain performance and effectiveness when deployed across large fleets or scaled to different AV/RV platforms with varying hardware and software configurations.

A-2: Modification-Free Approaches

Modification-free approaches can be classified into: external location verification, internal sensor fusion, and computer vision-based location verification.

- *External location verification*: This method utilizes existing GNSS signals (such as those from Galileo, GLONASS, and Beidou) and wireless network signals for location verification. While beneficial, this approach is partially foolproof, as these signals can be jammed or spoofed. For instance, Filppou et. al. [94] proposed a machine learning approach for detecting GPS location spoofing as anomalies. Specifically, the autonomous vehicle (AV) uses both V2V/V2I communication and GPS signals to compute its current location. A machine learning (ML) model is then trained using location data obtained from both V2V/V2I communication and attack-free GPS signals. Based on a predefined threshold, when the difference between these two locations exceeds the threshold, the ML model detects an anomaly. Similarly, Shabbir et al. [95] proposed a deep learning approach to detect GPS spoofing with the aid of signals from wireless communication infrastructure.
- *Internal sensor fusion*: This technique employs inertial measurement unit (IMU) sensors, including accelerometers, gyroscopes, and magnetometers, to corroborate GPS data like the work presented in [96]–[98]. However, it is prone to accumulative sensor errors over time, which can diminish its accuracy.
- *Computer vision-based location verification*: Utilizing computer vision to cross-reference physical-world landmarks and street signs with digital maps offers a promising solution. This method primarily requires software-level enhancements. For example, the work presented in [99] proposed an enhanced design of a computer vision-based localization system that can be integrated into AV navigation facilities to improve localization and mapping accuracy in extreme environmental conditions where GPS signals are unstable or absent. Based on our empirical study, the following conclusions can be drawn: A CNN-based ML model architecture, specifically YOLOv4, demonstrated superior performance compared to the two other investigated architectures and can be effectively employed as an image processor for AV localization applications. However, the reliability of these vision-based methods remains under scrutiny due to their vulnerability to minor adversarial manipulations, which can deceive image classifiers.

B. COUNTERMEASURE AGAINST LIDAR SPOOFING

In this subsection, we discuss the countermeasures against laser projection and object placement LiDAR spoofing attacks.

B-1: Countermeasures Against Laser Project Attacks

Several research studies have proposed detection methods to counter laser-projected LiDAR spoofing attacks in autonomous vehicles (AVs). In this subsection, we focus primarily on the most efficient countermeasures, such as those proposed in [50] and [16], and discuss their limitations.

- *Fake Shadow Detection (FSD)*: This methodology extends the work of [100] for detecting Physical Removal Attacks (PRA), by identifying shadow regions in the point cloud within the Region of Interest (ROI) to detect potential ORA-based attacks. The FSD approach utilizes projection and frustum techniques to filter out shadows, flagging those larger than a threshold (15 cubic meters) as removal attacks. When evaluated, it achieved a True Negative Rate (TNR) of 82.5% and a True Positive Rate (TPR) of 91.2%. However, analyzing

TABLE 6. LiDAR Spoofing Countermeasure.

Paper	Cost	Deployment Overhead	Effectiveness	Scalability	Summary and Practical Feasibility
[16]	Primarily software-based, which is less costly than hardware alternatives.	Significant changes to existing LiDAR-based perception systems are required.	Addresses specific vulnerabilities (Occlusion Patterns, Data Density, and Distance), effectively.	Adaptable to existing systems, though dependent on system architecture.	Effective software-based solution with low cost but potentially high deployment overhead. **Dataset:** KITTI [75].
[50]	Software-centric approach suggests cost savings over hardware solutions.	Needs considerable alterations to current LiDAR-based perception systems.	Tailored to counteract LiDAR-specific vulnerabilities.	Software adaptability ensures scalability, subject to system-specific architecture.	Cost-efficient and effective with a focus on software, but with moderate to high deployment overhead. **Dataset:** KITTI [75] and LGSVL [93].
[20]	Increased compute cost due to training with adversarial examples.	Requires recalibration for different environments, which is a noted limitation.	Effectiveness is environment-specific and inconsistent.	Recalibration needs could limit application across varying conditions.	Provides moderate cost countermeasures with environment-specific effectiveness and limited scalability. **Dataset:** KITTI [75].
[11]	Requires 3D printing and pre-attack scene analysis but avoids high-end hardware modifications.	Needs adversarial object design, pre-attack environmental mapping, and deployment.	Successfully misleads both LiDAR and camera-based perception with over 90% success rate.	Attack success is affected by variations in scene lighting, object material, and environmental dynamics.	The study demonstrates an attack where adversarial 3D objects can fool both LiDAR and camera-based perception systems, leading to a complete failure in detecting obstacles. It challenges the assumption that multi-sensor fusion enhances security, highlighting the need for improved defenses against shape manipulation attacks. It requires high end hardware modifications. **Dataset:** KITTI [75].

shadow patterns introduces computational complexity, which may cause delays in real-time decision-making systems, such as those used in autonomous vehicles.

- *Azimuth-based Detection*: This method is designed to detect removal attacks by analyzing disparities in the raw point cloud data, specifically within the horizontal angular view (azimuth) of the LiDAR. An attack is indicated by a gap exceeding a 1-degree angle in azimuth values. Azimuth-based detection sorts cloud points by their azimuth values to identify such gaps. It achieved a 99.98% TNR on the KITTI dataset and a 100% TPR on synthesized and real-world attack scenes. Furthermore, it effectively revealed the attack angle and direction in these scenarios. However, despite its high effectiveness against PRA, Azimuth-based Detection is not applicable to adversarial spoofing or other ORA-based attacks. On a high-performance system, the method demonstrated an average runtime of 7.9 ms per scene.
- *CARLO (oCclusion-Aware hieRarchy anomaLy detectiOn)*: CARLO functions as a practical post-detection module that eliminates the need for re-training models, a process that can be resource-intensive. It is also realistic in its approach, as it does not rely on white-box access to the model, ensuring broader applicability. This model-agnostic defense leverages occlusion patterns as invariant physical features to detect spoofed fake vehicles. CARLO employs two key components: free space detection and laser penetration detection, which utilize inter- and intra-occlusions to effectively identify spoofed vehicles.
- *Sequential View Fusion (SVF)*: SVF is a general architecture designed to enhance the robustness of LiDAR-based perception by embedding physical information into model learning, which necessitates re-training. This method incorporates the front-view representation of LiDAR point clouds into the detection model. By employing a semantic segmentation module, SVF effectively leverages front-view features, significantly reducing attack success rates while maintaining original performance. Although SVF outperforms CARLO in

defense performance, it incurs a slight decrease in Average Precision (AP), indicating the need for additional training efforts.

B-2: Countermeasures Against Shape Manipulation Attacks

LiDAR-camera fusion and shape manipulation attacks pose a significant threat to the integrity of perception in AVs. It has long been believed that combining various sensory inputs through MSF improves resistance, but recent studies (e.g. [11]) argue differently. Experiments demonstrate that adversarially manipulated objects can be crafted to fool both LiDAR and camera systems at the same time, revealing vulnerabilities within sensor fusion algorithms. By altering an object's perceived position and shape, attackers can take advantage of discrepancies in fusion models to cause misclassification or a complete failure to detect relevant obstacles. Numerous strategies, including multi-sensor redundancy techniques, adversarial object detection methods, and scene verification mechanisms, have been suggested by researchers as countermeasures. However, each of these approaches has its own limitations, which underscores the need for more adaptable and robust AV perception systems required.

- *Fusion-Based Multi-Sensor Defense*: MSF systems try to cross-validate sensor observations leveraging LiDAR point cloud geometric consistency as well as camera textural features. Combining depth and geospatial information, MSF systems can detect structural irregularities typical of manipulated attacks.
- *Adversarial Object Detection*: Advanced detection models leverage point cloud perturbation analysis and spatial inconsistency detection to mark adversarially manipulated objects. Object distortions introduced artificially are removed by filtering techniques such as statistical outlier removal and clustering-based outlier rejection.
- *Physical Scene Verification*: The approach entails matching current perception output with stored environmental knowledge for object integrity verification. Reference objects such as map-based priors and past sensor data assist in distinguishing real objects from distorted objects.

B-3: Countermeasures Against Object Placement Attacks

One of the efficient countermeasures against object placement attacks on AVs is the approach proposed in [20], which can be categorized into the following methods:

- *Adversarial Location Defense*: This method involves modifying LiDAR model training to minimize adversarial scores near vehicle edges, making it more challenging for attackers to exploit these areas. However, this approach may impact detection accuracy as it seeks to balance defense effectiveness.
- *Adversarial Training*: Incorporating adversarial examples into the training data helps reduce the success rate of attacks, though some attacks may remain partially effective.
- *Sensor Fusion*: This approach combines LiDAR with other sensors, such as cameras and radar, to achieve more robust object detection. However, it may lead to increased system cost and complexity.

Table 5 provides a summary of some of the LiDAR spoofing countermeasures in terms of cost, deployment overhead, effectiveness, and scalability.

C. COUNTERMEASURE AGAINST CAMERA SPOOFING

This subsection discusses the sticker pasting and light projection camera attacks.

C-1: Countermeasures Against Sticker Pasting Attacks

We focus on machine learning model-level defense strategies for automated lane centering (ALC) systems, as explored in [22]. These strategies include:

- *Model Input Transformation Method*: This includes several methods, such as JPEG compression, bit-depth reduction, adding Gaussian noise, median blurring, and autoencoder reformation to the input.

C-2: Countermeasures Against Light Projection Attacks

The research in [33] evaluates three defense mechanisms against physical adversarial examples (AE):

- *SentiNet*: A defense specifically designed to detect physical AEs. SentiNet uses saliency masks to identify suspect regions in images. However, the SLAP (Short-Lived Adversarial Perturbations) AEs effectively bypassed SentiNet detection in over 95% of evaluated frames.
- *Input Randomization*: This approach involves randomizing input image parameters. The study found that input randomization was unable to detect the SLAP AE. Moreover, this method unexpectedly reduced the model's accuracy, indicating that the original models needed to be trained with sufficient data augmentation.
- *Adversarial Learning*: This involves retraining models with an added FGSM-adversarial loss to make them more resistant to adversarial attacks. While adversarial trained models showed a slight accuracy decrease on test sets, this method proved to be a more suitable defense against SLAP, significantly reducing the attack success rate.

Table 6 provides a summary of some of the Camera spoofing countermeasures in terms of cost, deployment overhead, effectiveness, and scalability.

IV. AVs ATTACKS SIMULATION FRAMEWORKS

In this section, we discuss the most commonly used simulation environments and frameworks in the literature for simulating attacks and detection mechanisms in AVs.

TABLE 7. Camera Spoofing Countermeasure.

Paper	Cost	Deployment Overhead	Effectiveness	Scalability	Summary and Practical Feasibility
[22]	Requires additional computational resources for model-level defenses but does not require hardware modifications.	No changes to existing vehicle hardware but requires modifications in lane detection models.	Model input transformations such as JPEG compression, bit-depth reduction, Gaussian noise, median blurring, and autoencoder reformation were tested, but all methods resulted in trade-offs between attack resistance and lane detection accuracy.	Transferability to other Automated Lane Centering (ALC) systems remains uncertain due to limited access to production ALC models.	Evaluates lane-detection model defenses against sticker-based attacks and highlights the need for model-specific adversarial training rather than general-purpose input transformations. **Dataset:** comma2k19 [76].
[33]	Requires projector-based adversarial perturbations, avoiding permanent physical modifications.	Attack success depends on the environment (lighting conditions, projection angles) and requires precise calibration.	Effective against multiple detection models, bypassing SentiNet-based defenses with a success rate of over 95%. However, adversarial training can mitigate its effectiveness.	The attack's success is constrained by ambient lighting and projection surface material.	Demonstrates a novel projection-based adversarial attack (SLAP) that manipulates object perception using a projector. Highlights the limitations of traditional physical AE defenses and suggests adversarial training as a countermeasure. Its practicality relies on the environment conditions and calibration precision. **Dataset:** GTSRB [77].

A. ADoPT FRAMEWORK

Anomaly Detection based on Point-Level Temporal Consistency (ADoPT) is used to identify unusual or unexpected behaviors in data over time, particularly when data points are connected by temporal relationships. The framework aims to detect anomalies by examining the consistency of individual data points in a temporal sequence or time series. For AVs, ADoPT can be applied to detect abnormal navigational perception sensor readings or unusual behavior in vehicle systems. ADoPT was introduced in [101], focusing on detecting LiDAR spoofing attacks on AVs. Specifically, ADoPT measures the temporal consistency at the point cloud level and performs better than existing methods such as CARLO and 3D-TC2. It achieves these results by delivering lower false positive rates and higher true positive rates. Currently, ADoPT focuses on single-frame fake object injection attacks; however, it also has the ability to address LiDAR spoofing attacks that span consecutive frames. ADoPT achieves this by promptly eliminating detected spoofing points, converting them into benign frames, and continuously identifying spoofed objects in subsequent incoming frames. Cho et al. [101] mention that future research with ADoPT will involve evolving it to counter various LiDAR data manipulation attacks, such as removal attacks, thus enhancing the overall robustness of the perception modules within the AVs.

B. NADs FRAMEWORK

Network Anomaly Detection Systems (NADs) were originally designed to detect unusual patterns, behaviors, or activities in a computer network that may indicate malicious activity, faults, or security threats. In the context of AVs, Meyer et al. [102] developed a comprehensive assessment framework for NADs and explored their susceptibility to cyberattacks. The authors argue that the security and safety of future AVs depend upon the holistic protection of automative components, particularly the onboard Time Sensitive Networks (TSNs). These TSNs require monitoring for both safety and security; however, a thorough evaluation of anomaly detection methods in this specific context requires further research. The proposed NADs framework allows for reproducible, comparable, and rapid evaluation of detection algorithms. It is also based on a simulation toolchain that contributes configurable topologies, traffic streams, anomalies, attacks, and detectors. The key contributions of [102] include

utilizing a comprehensive simulation environment with individual configurations, automatic labeling, and replaceable attack models to generate datasets. These datasets consist of labeled Packet Capture (PCAP) files that describe various scenarios, including both contain benign and abnormal communication. The NADs framework allows for the evaluation and comparison of different combinations of traffic filters, metrics, and algorithms. NADs are crucial for securing future automotive systems and mitigating risks. Furthermore, the authors highlight that the performance of NADs relies on the TSN traffic class, the type of anomaly, and the specific implementation details of the NADs. The NADs framework is adaptable, extendable, and proves suitable for both fast and insightful examinations.

C. RoboTack FRAMEWORK

RoboTack, introduced in [103], is an advanced form of malware strategically designed to compromise the perception systems of AVs, posing a significant security threat. Unlike random attacks, which involve arbitrary miscalculations of trajectories for randomly chosen non-autonomous vehicles or pedestrians at random times and durations, RoboTack demonstrates a much higher efficacy in creating safety hazards. In driving simulations, random attacks trigger emergency braking in only 2.3% scenarios. In contrast, RoboTack RoboTack forced emergency braking in approximately 75.3% of the simulations, making it 33 times more successful in creating safety hazards. Furthermore, while random attacks did not result in accidents, RoboTack was responsible for causing accidents in the majority of its runs, particularly in scenarios involving pedestrians. Interestingly, RoboTack was less successful in scenarios involving other vehicles. This discrepancy is attributed to Apollo's[1] perception system, which the study identified as being less effective at detecting pedestrians compared to vehicles. Specifically, RoboTack required only 14 camera frames featuring pedestrians to cause an accident, whereas it needed 48 consecutive frames involving vehicles to achieve the same outcome. The primary goal of the work introduced in [103] was to underscore the critical need for robust security in AV systems. By demonstrating how a malicious attacker can efficiently exploit vulnerabilities to target AV safety and cause catastrophic outcomes, their work highlights the importance of understanding the how, when, and what of potential attacks. Additionally, their findings serve as a stark warning about the vulnerabilities in autonomous vehicle systems and emphasize the urgent need for comprehensive security measures to protect AV technologies.

D. SADC FRAMEWORK

Sensor Attack Detection and Classification (SADC) framework, introduced by Begum et al. [104], is designed to simulate and detect attacks on AVs in a sixth-generation (6G) vehicle-to-everything (V2X) (6G-V2X) environment with attack detection. The framework addresses emerging cyberattacks by providing an effective mechanism for sensor attack detection and classification in AVs. It is capable of mitigating various modern assaults, including replay attacks, False Data Injection (FDI) attacks, and stealthy attacks, thereby reducing the risk of catastrophic accidents. The SADC framework incorporates GPS and LiDAR sensor attack detectors alongside a Pattern-based Attack Classification (PAC) algorithm. The GPS and LiDAR detectors analyze discrepancies between multiple sensor measurements to identify abnormal behaviors indicative of sensor attacks. Based on this analysis, a protocol-based attack detection scheme identifies compromised source sensors. The PAC algorithm further classifies malicious AVs into distinct categories based on patterns of abnormal sensor behavior [104]. In summary, the SADC framework leverages a modular 6G-V2X network, enabling AVs to collaboratively detect cyberattacks by sending alarm signals. This results in rapid and accurate attack detection, significantly reducing the likelihood of accidents. Furthermore, the framework demonstrates superior accuracy compared to other approaches in both detecting and classifying sensor attacks effectively.

E. SIMUTACK FRAMEWORK

Simutack is an open-source attack simulation framework for connected autonomous vehicles (CAVs), introduced in [105]. SimuTack provides a safe and realistic environment for conducting extensive security tests and generating reliable vehicular data under a variety of attack conditions. It covers a broad range of automotive attack targets and includes sample implementations of well-known attack vectors, such as sensor attacks, network attacks, and V2X attacks. These attack vectors highlight significant vulnerabilities in CAV systems and are demonstrated in SimuTack through three realistic use cases: a GNSS attack, a camera blinding attack, and a V2X attack. These scenarios were specifically chosen for their relevance in showcasing the practicality and effectiveness of SimuTack in a realistic driving environment. SimuTack is built upon CARLA, a widely used simulation tool in both academia and industry. CARLA provides simulated sensor data, ensuring that the generated data is both realistic and trustworthy. Additionally, Finkenzeller et al. [105] proposed a comprehensive attack layer in SimuTack, enabling the simulation of highly relevant and complex attacks against CAV systems. By incorporating this attack layer, SimuTack enhances existing automotive testing environments, such as CARLA, by enabling the design and testing of secure autonomous systems. This makes it a valuable tool for advancing the security and reliability of connected and autonomous vehicles.

Table IV-B provides a summary of the simulation platforms proposed in the literature discussed in this section.

[1]Apollo is an open-source platform for autonomous driving, developed by Baidu company. It serves as a comprehensive framework for building, testing, and deploying AVs.

TABLE 8. Literature Comparison: Simulation Environments, Frameworks, Attacks, and Detection Mechanisms in AVs.

Paper	Simulator	Framework/Defenses	Attacks	Open-Source
[101]	None (evaluated on nuScenes dataset, no specific simulator)	ADoPT Framework	LiDAR Spoofing Attack	No
[102]	Custom simulation toolchain (no mainstream simulator like CARLA or Apollo)	NADs	Anomalies	Yes
[103]	Robotack	Robotack	Smart Malware	No
[104]	Not specified (no known mainstream simulator)	SADC	False Data Injection, Stealthy Attack, Replay Attack	No
[105]	CARLA	Simutack	GNSS Attack, Camera Binding Attack, V2X Attack	Yes
[106]	Custom hardware-in-the-loop system (no CARLA or Apollo)	EMI-LiDAR	Electromagnetic Interference (EMI)	No
[107]	None (real-world + testbed, no CARLA or Apollo)	MSF-ADV	Physical Attacks	Yes
[108]	None (no mainstream simulator used)	R2L-SLAM	Poor Sensor Conditions	No
[109]	CARLA	GPS-IDS	GPS Spoofing Attack	Yes
[110]	None (real sensors in controlled experiments)	SoundFence	Signal Injection Attack	No
[111]	SVL Simulator	Test Generation Technique	Failure Revealing Scenarios	Yes
[112]	Not specified clearly (based on datasets and real world)	Self-Localization Framework	Localization Drift	No
[113]	ROS-based (no dedicated simulator like CARLA or Apollo)	ROS	Human Error in AVs	Yes
[114]	CARLA	CARLA Autoware Bridge	Sensor Attacks, Vehicle Accidents	Yes

F. CARLA AND Its EXTENSIONS

CARLA is an open-source simulator developed from scratch to support the development, training, and validation of autonomous urban driving systems. It offers significant advantages for AV research, such as generating reliable sensor data and facilitating efficient system-level testing. However, CARLA also presents challenges, including the difficulty of achieving seamless compatibility between its software stack and simulation environment. Additionally, its specific design limits its applicability to other AV software or simulators.

Several simulation frameworks have been built upon CARLA, aiming to address some of these challenges and enhance its functionality, particularly for attack simulations. For instance, SimuTack [105] leverages CARLA to generate trustworthy sensor data and incorporates a comprehensive attack layer for simulating relevant attacks on AVs. This enhancement allows for the safe evaluation of individual component security by generating desired attack scenarios and analyzing their effects. Despite these improvements, SimuTack has its own limitations. Like CARLA, it faces challenges with seamless software stack integration, and its specific design for CARLA restricts its usability with other AV software or simulators.

In another development, Kaljavesi et al. [114] introduced a bridge that connects the CARLA simulator with Autoware Core or Autoware Universe, representing a significant advancement in the field of autonomous driving modules. This open-source bridge addresses the compatibility challenge by enabling efficient system-level testing of autonomous driving modules. It provides researchers with a user-friendly platform to test their modules within an integrated software framework. Furthermore, it establishes a reliable communication system, enhancing the overall functionality of the autonomous driving modules.

Despite these benefits, the bridge faces its own challenges. Its implementation is complex, and like SimuTack, its design is tailored to CARLA and Autoware, limiting its broader applicability to other AV software or simulators.

G. BAIDU APOLLO AND Its EXTENSIONS

Baidu Apollo [116] is an open-source platform developed by Baidu, designed to deliver intelligent vehicle solutions and advance autonomous driving technologies.

Several simulation frameworks and studies have leveraged Baidu Apollo to address challenges and extend its capabilities, particularly for testing and attack simulation. Ebadi et

TABLE 9. Limitations of simulation frameworks currently used in the literature.

Framework	Fidelity Limitations	Simulation Realism	Sensor Modeling Limitation
ADoPT [101]	Focus solely on detecting LiDAR spoofing attacks on AVs.	The framework lacks of various LiDAR data manipulation attacks and removal attacks.	Relies on the nuScenes dataset [115] and lacks detailed sensor modeling, such as varying noise levels.
NADs [102]	Neglects latencies caused by the interaction between sensors.	Focus only on In-Vehicle Network (IVN) attacks that lead to operational disruptions and loss of control to safety-critical accident.	Does not consider sensor modeling, but use existing datasets.
RoboTack [103]	Neglects hardware imperfections in mmWave radar and lacks severe weather variation.	Realistic traffic flow with Monte Carlo sampling, but urban complexity remains a challenge.	Radar simulation uses basic models that does not accurately represent complex real-world interactions, such as multipath effects and material-dependent reflections.
SADC [104]	Focus on false data injection and detection, excluding the simulation of the attack itself.	Uses Simulation of Urban MObility (SUMO) to simulate to support urban and highway traffic scenarios.	GPS and LiDAR sensors are not modeled, but they use existing datasets to false data injection attacks.
Simutack [105]	Difficulty capturing complex attack surfaces and sensor imperfections.	Focus on cybersecurity attacks with realistic vehicle and environment simulation.	Relies on CARLA and SUMO, so it does not consider the noise profiles and imperfections in sensor modeling.
CARLA extension [114]	Does not capture complex real-world vehicle dynamic, such as suspension nuances or extreme weather effects.	Consider day/night cycles, some weather effects .	Doesn't always match the noise profiles, occlusions, or hardware imperfections seen in real sensors.
Baidu Apollo and extensions [111]–[113], [116]	Utilize rule-based vehicle dynamics models, making it difficult to capture nonlinear vehicle behaviors, such as high-speed maneuvers or adverse weather scenarios.	Integration with CARLA and SVL enables scalable testing.	Does not consider sensor fusion challenges, such as synchronization issues, varying sensor noise characteristics, and environmental interferences.

al. [111] integrate the Baidu Apollo platform with the SVL simulator to evaluate pedestrian detection and emergency braking systems using an evolutionary automated test generation technique. This method effectively creates failure-revealing test cases, serving as a powerful tool for quality assurance. However, the vast space of test input parameters presents a challenge, making efficient scenario generation complex.

Xia et al. [112] propose a GPS-independent self-localization framework for autonomous vehicles, tested on the Baidu Apollo Southbay dataset. This lightweight framework is optimized for embedded systems, balancing hardware and software constraints. While its design ensures efficient operation in resource-limited environments, its reliance on specific hardware and software configurations may limit its versatility across diverse applications.

Raju et al. [113] examine ROS-based open software platforms, including Autoware and Apollo, highlighting their role in simplifying the development of autonomous vehicle platforms. These frameworks eliminate the need to build systems from scratch by providing robust libraries and algorithms. However, their reliance on preexisting components may restrict customization and optimization for specific autonomous driving scenarios.

H. CARLA AND BAIDU APOLLO COMPARISON

Both CARLA and Baidu Apollo offer distinct advantages in the field of AV research. CARLA excels in generating reliable sensor data and enabling system-level testing, though its design can limit compatibility with other software or simulators [105], [114]. On the other hand, Baidu Apollo has demonstrated effectiveness in testing critical systems, such as pedestrian detection and emergency braking, and supports a GPS-independent self-localization framework [111]–[113]. However, Apollo's reliance on specific hardware and software configurations may hinder its versatility in diverse applications.

Both platforms are ROS-based, which simplifies the development of autonomous vehicle systems by providing robust libraries and tools. Nevertheless, their dependence on preexisting components can limit adaptability and optimization for specific scenarios.

Based on our literature survey, Baidu Apollo emerges as more prevalent in advancing autonomous vehicle technology, particularly for its comprehensive capabilities in addressing real-world AV challenges.

Table IV-F provides a summary of the limitations of simulation frameworks currently used in the literature. Specifically, the table address fidelity limitations, sensor modeling accuracy, and realism in attack/defense evaluation, thereby helping researchers identify suitable platforms and understand current constraints in AV security experimentation. Regarding hardware testbeds, which rely on physical components, metrics such as simulation realism and sensor modeling accuracy are not applicable.

In this section, we explore the testbed environments currently used in the literature for implementing attacks and detection mechanisms in AVs.

I. EMI-LIDAR

Electro-Magnetic-Interference (EMI) LiDAR testbed, introduced in [106], reveals a new vulnerability in LiDAR sensors and Time-of-Flight (ToF) circuits in AVs that affects object detection algorithms. This study demonstrates that EMI can disrupt AV object detection models, such as PointPillars [117], PointRCNN [118], and Apollo [116]. Moreover, it uncovers a new threat known as Intentional EMI (IMEI). For this experiment, the authors use several components of a ToF circuit used in a LiDAR system for AVs. The setup consists of a LiDAR, an antenna, a DC circuit, a Universal Software Radio Peripheral (USRP), and an amplifier, all housed within a Faraday Cage. According to their research, IEMI affects the ToF circuits of modern LiDARs, and these vulnerabilities can be exploited to cause the AV perception system to misdetect and misclassify objects, as well as perceive non-existent obstacles. Furthermore, these vulnerabilities are evaluated in three AV perception modules, showing a classification rate drop below 50%. The testbed specifically targets spinning LiDAR sensors used by companies like General Motors' Cruise [119] and Google's Waymo. The authors also propose two strategies to detect signal injections: Ground Point-Based Detection and Point Intensity-Based Detection. Ground Point-Based Detection extracts the ground points from the LiDAR scene (e.g., using a ground point segmentation model integrated with Apollo-based perception systems) and analyzes them to detect potential IEMI injection. Point Intensity-Based Detection uses a similar methodology, where potential IEMI injection is detected if a region's average point intensity is below a given threshold.

J. MSF-ADV

MSF based Autonomous Driving (AD) perception, introduced in [107], integrates AD systems deployed in Level 4 AVs and MSF design. It focuses on implementing attacks on in-road obstacle detection systems that use camera and LiDAR fusion. The attacks challenge the basic MSF design assumption in the AD context by successfully engineering a physical-world adversarial attack. This attack creates adversarial 3D objects that mislead a victim AV, causing a failure in detection and potentially leading to a crash. The authors use an optimization-based approach to address two main design challenges. The first challenge is non-differentiable target camera and LiDAR sensing systems, and the second is non-differentiable cell-level aggregated features used by LiDAR. For this experiment, they used an iPhone 8 Plus back camera and a Velodyne VLP-16 LiDAR to collect images and point clouds. Adversarial traffic cone meshes are generated and printed at a 1:6:67 scale with 380 um precision, simulating a scenario of sensors installed on a car driving on a standard 3.6-meter-wide highway.

To enhance the attack's robustness, stealthiness, and physical world realizability, they implement specific strategies. The attack is evaluated on MSF algorithms included in representative open-source industry-grade AD systems (such as Baidu Apollo and AutoWare.AI [120]) in real-world driving scenarios. It achieves a 91% success rate across different object types and MSF algorithms. Furthermore, the adversarial attack demonstrates high stealthiness, robustness to victim positions, transferability across MSF algorithms, and physical-world realizability after being 3D-printed and captured by LiDAR and camera devices. The authors also apply the proposed attack to a production-grade simulator, showing that it achieves a 100% vehicle collision rate with an industry-grade AD system. They also evaluate and discuss defensive strategies, such as Bit-depth reduction, median smoothing, JPEG compression, and Autoencoder reformation.

K. R2L-SLAM

Radar-to-LiDAR Simultaneous Localization and Mapping (R2L-SLAM), introduced in [108], is a testbed that incorporates mmWave radar into the sensing setup for a 2D LiDAR SLAM application. This testbed addresses the reliability issues of optical sensing modalities in AVs under adverse environmental conditions (e.g., rain or smoke). Such conditions pose significant safety concerns, as optical sensors may fail to accurately perceive obstacles. To mitigate these concerns, the authors propose augmenting AV perception suites with additional sensing modalities alongside LiDAR sensors. Specifically, they introduce a modality prediction method by integrating a single-chip mmWave radar sensor with an existing 2D LiDAR setup. To facilitate this, the researchers created a custom dataset for training and validation, which includes time-synchronized, raw mmWave radar data and LiDAR measurements. The testbed is constructed using a Texas Instruments IWR1443 single-chip mmWave radar sensor and a Hokuyo UST-20L LiDAR sensor [121], mounted on a small robotic platform. The training dataset comprises approximately 10,000 synchronized samples, while the validation dataset contains 1,500 samples. Due to the inherent differences in sensing modalities, mmWave radar provides a sparser environmental representation compared to LiDAR. The core innovation lies in the R2L-Net model, which upscales the sparse radar point-cloud data into a dense, LiDAR-like point-cloud. This enhancement significantly improves the AV's perception capabilities under challenging environmental conditions, such as smoke and fog, thereby enhancing safety and reliability.

TABLE 10. Summary of the testbed environments currently used in the literature for implementing attacks and detection mechanisms in AV, along with their limitations.

Testbed	Hardware Used	Attacks	Defense	Limitations
EMI-LiDAR [106]	ToF Circuit in LiDAR System	EMI is used to compromise LiDAR's ToF circuits	EMI-LiDAR	LiDAR systems typically exhibit resistance to EMI due to rigorous Electromagnetic Compatibility (EMC) testing and the incorporation of anti-interference designs, such as shielding and layout optimization, which mitigate EMI effects [122].
MSF-ADV [107]	iPhone 8 Plus Camera and Velodyne VLP-16 LiDAR	Physical-World Attacks	MSF-ADV	The attack targets a specific scene or road section, making it non-universal and scene-dependent.
R2L-SLAM [108]	2D LiDAR Sensor	Poor Sensor Conditions	R2L-Net	LiDAR and radar modalities are treated separately, without allowing for the fusion or joint reasoning of their predictions. Additionally, the approach requires a large, well-synchronized dataset across all sensor modalities, which limits scalability and generalization.
GPS-IDS [109]	Autonomous Rover, Raspberry Pi-4 model B and HackRF One SDR	GPS Spoofing Attack	GPS-IDS	The dataset suffers from class imbalance, with significantly more normal samples than malicious ones, which can bias models during training.
SoundFence [110]	Arduino UNO Rev3 and HC-SR04	Signal Injection	SoundFence	Lacks analysis of full environment effects, which can significantly impact the performance and accuracy.

L. GPS-IDS

The GPS intrusion detection system (GPS-IDS), introduced in [109], is an anomaly behavior analysis-based intrusion detection framework designed to detect GPS spoofing attacks on AVs. This framework employs a unique physics-based vehicle behavior model, integrating a GPS navigation model with the conventional dynamic bicycle model to accurately represent AV behavior. The model captures lateral dynamics and vehicle states to provide a precise representation of normal AV navigation behavior. Temporal features derived from this model are used to differentiate normal behavior from attacks using machine learning techniques. For this experiment, the authors developed an autonomous vehicle testbed. The hardware architecture comprises a Rover Unit and an Attacker Unit. The Rover Unit includes a one-tenth scale radio-controlled truck chassis, a Pixhawk flight controller, a Neo M8N GPS receiver with a built-in compass, telemetry modules, and radio transmitter-receiver modules. The Attacker Unit uses a Raspberry Pi-4 Model B to generate spoofed GPS baseband signal data streams. These streams are converted to radio frequency by a HackRF One software-defined radio (SDR) and transmitted using an antenna to spoof GPS signals. The proposed GPS-IDS framework was evaluated using the AV-GPS-Dataset and achieved an F1 score of up to 94.4%. Compared to the Extended Kalman Filter (EKF) detector implemented in the experimental setup, the framework demonstrated a detection time improvement of up to 13 seconds. These results highlight the effectiveness of GPS-IDS in detecting GPS spoofing attacks and enhancing the security of AVs.

M. SOUNDFENCE

SoundFence, proposed in [110], is a physical-layer defense system that leverages the signal processing capabilities of sensors without requiring additional hardware. The authors provide a comprehensive analysis of the vulnerabilities of ultrasonic sensors in AVs to signal injection attacks. These attacks involve injecting malicious acoustic signals to create fake obstacles, potentially leading to incorrect AV decision-making and disastrous consequences.

The proposed defense system operates directly on commercial ultrasonic sensors, enabling it to reject malicious signals or sensor readings by effectively detecting abnormal sensor behavior with minimal false alarms. Furthermore, it can accurately distinguish real echoes from injected signals, thereby identifying and mitigating injection attacks.

In their study, the authors systematically analyzed various attack models, considering attackers with differing levels of sophistication. They found that even low-powered adversaries could consistently execute signal injection attacks under certain conditions, creating persistent fake obstacles with relative ease. These findings highlight the critical threat posed by signal injection attacks and emphasize the need to address these vulnerabilities in AV systems' design and assembly.

Finally, the paper underscores the importance of secure integration between the cyber and physical components of AV systems to ensure safe and reliable operation. This research offers valuable insights for designing and evaluating the security of AV sensing systems, emphasizing the need for robust defense mechanisms in autonomous and semi-autonomous vehicles

Table 10 provides a summary of the testbed environments

discussed in this section.

V. KEY INSIGHTS ON FEASIBILITY OF ATTACKS ON OFF-ROAD (NON-URBAN) ENVIRONMENT

The analysis above has provided several critical insights and learning scenarios, which are pivotal in advancing our understanding of AV technology. In this section, we aim to offer a first-order feasibility analysis to determine whether these attacks can be carried out in off-road settings. Specifically, we seek to answer a simple question: can these attacks be readily deployed in off-road environments? The short answer is yes; many of the attacks reviewed can indeed be deployed in off-road settings. However, for a more nuanced analysis, we categorize these attack methodologies into two distinct groups:

1) Attacks requiring a comprehensive hardware setup: These attacks demand bulky equipment, multiple hardware devices, and/or present challenges to implementation due to their complexity.
2) Attacks executed without a full hardware configuration, such as laser and white-box attacks.

A. ATTACKS REQUIRING A COMPREHENSIVE HARDWARE SETUP

We provided an analysis for each type of sensor in the following:

- **GPS Spoofing [9]–[13]**: This is a complex attack that requires a comprehensive hardware setup to generate false GPS signals, making it difficult to implement in off-road scenarios. While it can be deployed in off-road settings, the cost and complexity of the attack setup pose significant challenges for its practical use in such environments.
- **IMU Spoofing [40]–[44]**: This attack uses acoustic signals to manipulate IMU data, corrupting the motion detection signals in AVs and UAVs. Performing this attack requires an expensive and complex hardware setup, making IMU spoofing difficult to execute. Similar to GPS spoofing, these attacks can be deployed in off-road scenarios, but the high cost and complexity of the hardware setup make them challenging to implement in such environments.
- **Ultrasonic Sensor Spoofing [15], [47]**: Creating an ultrasonic pulse attacker requires an expensive and comprehensive hardware setup, complicating ultrasonic spoofing attacks. Although these attacks can be executed in off-road scenarios, the costly hardware setup makes deployment challenging.
- **mmWave Radar Spoofing [48], [49]**: These attacks can be either passive or active. Passive attacks require access to the victim's objects, while active attacks require hardware devices (e.g., USRP SDR) to replay the radar signal. Such attacks can be deployed in off-road scenarios, but passive attacks are often impractical and easy to detect by the human eye. In contrast, active attacks require an expensive hardware setup to create spoofed signals, making them challenging to deploy.

B. ATTACKS EXECUTED WITHOUT A FULL HARDWARE CONFIGURATION

We reviewed attacks that are executed without full hardware in the following scenarios:

- **LiDAR Spoofing [11], [16]–[18], [20]**: Light projection requires only a laser gun, making it relatively easy to implement. Another type of LiDAR attack, shape manipulation, alters the shape of an object to deceive LiDAR and camera sensors. The main challenge for this attack is whether the attacker has knowledge of the off-road terrain where the AV is expected to operate. Object placement is another LiDAR spoofing-based white-box attack, where the attacker exploits the vulnerabilities of the LiDAR sensor and creates an adversarial object using 3D printing to deceive the sensor. The primary challenge here is whether the object can be discreetly placed during the AV's operation.
- **Camera Spoofing [27]–[30], [33]–[38]**: Camera spoofing using sticker pasting involves training a DNN on adversarial image examples (AEs). The effectiveness of these images in off-road environments still needs to be explored. However, in theory, if the attacker has some knowledge of the off-road environment, these AEs can be specifically designed, making the attack feasible.

VI. CONCLUSION AND FUTURE DIRECTIONS

This survey highlights the vulnerabilities of AVs' navigational perception sensors and the various attack vectors that exploit them. Specifically, it emphasizes attacks that disrupt AV's decision-making process, effecting both navigation and perception, which are crucial for secure and safe AV operation. While considerable progress has been made in understanding these threats and developing countermeasures, key challenges remain, especially in defending against sensor fusion attacks and validating defense mechanisms under real-world conditions. To bridge these gaps, we conducted an extensive review of the literature on attacks targeting various types of sensors, including GPS, LiDAR, radar, and camera systems, as well as multi-sensor fusion. We also analyzed a range of defense mechanisms and countermeasures, offering insights into strengthening AV cybersecurity by enhancing detection mechanisms, improving sensor resilience, and developing more robust mitigation strategies. Furthermore, we presented key insights into the characteristics, methods, and impacts of attacks on navigation perception sensors.

Looking ahead, future research directions include:

- Expanding investigations to broader AV attack surfaces, such as vulnerabilities in electronic control units (ECUs) and AI-based decision-making systems.
- Developing standardized datasets and simulation environments for adversarial testing and benchmarking AV defense strategies.

- Enhancing real-time resilience in multi-sensor fusion systems, focusing on detection and recovery from coordinated, cross-sensor attacks.
- Exploring red-teaming methodologies to rigorously evaluate AV system robustness under sophisticated and adaptive threats.
- Investigating the application of Zero Trust Architecture (ZTA) in AV perception systems to enable continuous authentication and strict access control.
- Designing and implementing trustworthy AI models with uncertainty quantification and adversarial robustness to ensure safe and ethical decision-making.

By addressing these research challenges, we can contribute to the development of more secure, intelligent, and trustworthy autonomous transportation systems.

.

REFERENCES

[1] Seyyed Salman Alavi, Mohammad Reza Mohammadi, Hamid Souri, Soroush Mohammadi Kalhori, Fereshteh Jannatifard, and Ghazal Sepahbodi. Personality, driving behavior and mental disorders factors as predictors of road traffic accidents based on logistic regression. Iranian journal of medical sciences, 42(1):24, 2017.

[2] Asad J Khattak, Numan Ahmad, Behram Wali, and Eric Dumbaugh. A taxonomy of driving errors and violations: Evidence from the naturalistic driving study. Accident Analysis & Prevention, 151:105873, 2021.

[3] Hooman Razavi, Mohammad Reza Jamali, Morvaridsadat Emsaki, Ali Ahmadi, and Mostafa Hajiaghei-Keshteli. Quantifying the financial impact of cyber security attacks on banks: A big data analytics approach. In Proc. of IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 2023.

[4] Hooman Razavi and Mohammad Reza Jamali. Large language models (LLM) for estimating the cost of cyber-attacks. In Proc. of The 11th International Symposium on Telecommunications (IST), 2024.

[5] Mansi Girdhar, Junho Hong, and John Moore. Cybersecurity of autonomous vehicles: A systematic literature review of adversarial attacks and defense models. IEEE Open Journal of Vehicular Technology, 4:417–437, 2023.

[6] Prinkle Sharma and James Gillanders. Cybersecurity and forensics in connected autonomous vehicles: A review of the state-of-the-art. IEEE Access, 10:108979–108996, 2022.

[7] Farahnaz Javidi Niroumand, Parisa Ansari Bonab, and Arman Sargolzaei. Security of connected and autonomous vehicles: A review of attacks and mitigation strategies. In proc. of IEEE SoutheastCon, 2024.

[8] Xiaoqiang Sun, F. Richard Yu, and Peng Zhang. A survey on cybersecurity of connected and autonomous vehicles (CAVs). IEEE Transactions on Intelligent Transportation Systems, 23(7):6240–6259, 2022.

[9] Harshad Sathaye, Martin Strohmeier, Vincent Lenders, and Aanjhan Ranganathan. An experimental study of GPS spoofing and takeover attacks on UAVs. In Proc. of 31st USENIX security symposium (USENIX security 22), 2022.

[10] Kexiong Curtis Zeng, Shinan Liu, Yuanchao Shu, Dong Wang, Haoyu Li, Yanzhi Dou, Gang Wang, and Yaling Yang. All your GPS are belong to us: Towards stealthy manipulation of road navigation systems. In Proc. of 27th USENIX security symposium (USENIX security 18), 2018.

[11] Yulong Cao, Ningfei Wang, Chaowei Xiao, Dawei Yang, Jin Fang, Ruigang Yang, Qi Alfred Chen, Mingyan Liu, and Bo Li. Invisible for both camera and lidar: Security of multi-sensor fusion based perception in autonomous driving under physical-world attacks. In Proc. of IEEE symposium on security and privacy (SP), 2021.

[12] Daojing He, Yinrong Qiao, Shiqing Chen, Xiao Du, Wenjie Chen, Sencun Zhu, and Mohsen Guizani. A friendly and low-cost technique for capturing non-cooperative civilian unmanned aerial vehicles. IEEE Network, 33(2):146–151, 2018.

[13] Daojing He, Hong Liu, Sammy Chan, and Mohsen Guizani. How to govern the non-cooperative amateur drones? IEEE Network, 33(3):184–189, 2019.

[14] Yulong Cao, S Hrushikesh Bhupathiraju, Pirouz Naghavi, Takeshi Sugawara, Z Morley Mao, and Sara Rampazzi. You can't see me: Physical removal attacks on lidar-based autonomous vehicles driving frameworks. In Proc. of 32nd USENIX Security Symposium (USENIX Security 23), 2023.

[15] Chen Yan, Wenyuan Xu, and Jianhao Liu. Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle. Def Con, 24(8):109, 2016.

[16] Jiachen Sun, Yulong Cao, Qi Alfred Chen, and Z Morley Mao. Towards robust liDAR-based perception in autonomous driving: General black-box adversarial sensor attack and countermeasures. In Proc. of 29th USENIX Security Symposium (USENIX Security 20), 2020.

[17] Zizhi Jin, Xiaoyu Ji, Yushi Cheng, Bo Yang, Chen Yan, and Wenyuan Xu. Pla-lidar: Physical laser attacks against lidar-based 3D object detection in autonomous vehicle. In Proc. of IEEE Symposium on Security and Privacy (SP), 2023.

[18] Yulong Cao, Chaowei Xiao, Dawei Yang, Jing Fang, Ruigang Yang, Mingyan Liu, and Bo Li. Adversarial objects against lidar-based autonomous driving systems. arXiv preprint arXiv:1907.05418, 2019.

[19] Sri Hrushikesh Varma Bhupathiraju, Jennifer Sheldon, Luke A Bauer, Vincent Bindschaedler, Takeshi Sugawara, and Sara Rampazzi. Emi-lidar: uncovering vulnerabilities of lidar sensors in autonomous driving setting using electromagnetic interference. In Proceedings of the 16th ACM Conference on Security and Privacy in Wireless and Mobile Networks, pages 329–340, 2023.

[20] Yi Zhu, Chenglin Miao, Tianhang Zheng, Foad Hajiaghajani, Lu Su, and Chunming Qiao. Can we use arbitrary objects to attack lidar perception in autonomous driving? In Proc. of the 2021 ACM SIGSAC Conference on Computer and Communications Security, 2021.

[21] Yanmao Man, Ming Li, and Ryan Gerdes. GhostImage: Remote perception attacks against camera-based image classification systems. In Proc. of the 23rd International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2020), 2020.

[22] Takami Sato, Junjie Shen, Ningfei Wang, Yunhan Jia, Xue Lin, and Qi Alfred Chen. Dirty road can attack: Security of deep learning based automated lane centering under Physical-World attack. In Proc. of the 30th USENIX security symposium (USENIX Security 21), 2021.

[23] Jiakai Wang, Aishan Liu, Zixin Yin, Shunchang Liu, Shiyu Tang, and Xianglong Liu. Dual attention suppression attack: Generate adversarial camouflage in physical world. In Proc. of the IEEE/CVF conference on computer vision and pattern recognition, 2021.

[24] Yunhan Jia Jia, Yantao Lu, Junjie Shen, Qi Alfred Chen, Hao Chen, Zhenyu Zhong, and Tao Wei Wei. Fooling detection alone is not enough: Adversarial attack against multiple object tracking. In Proc. of International Conference on Learning Representations (ICLR'20), 2020.

[25] Yue Zhao, Hong Zhu, Ruigang Liang, Qintao Shen, Shengzhi Zhang, and Kai Chen. Seeing isn't believing: Towards more robust adversarial attack against real world object detectors. In Proc. of ACM SIGSAC conference on computer and communications security, 2019.

[26] Pengfei Jing, Qiyi Tang, Yuefeng Du, Lei Xue, Xiapu Luo, Ting Wang, Sen Nie, and Shi Wu. Too good to be safe: Tricking lane detection in autonomous driving with crafted perturbations. In Proc. 30th USENIX Security Symposium (USENIX Security 21), 2021.

[27] Lifeng Huang, Chengying Gao, Yuyin Zhou, Cihang Xie, Alan L Yuille, Changqing Zou, and Ning Liu. Universal physical camouflage attacks on object detectors. In Proc. of the IEEE/CVF conference on computer vision and pattern recognition, 2020.

[28] Kevin Eykholt, Ivan Evtimov, Earlence Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, and Dawn Song. Robust physical-world attacks on deep learning visual classification. In Proc. of the IEEE conference on computer vision and pattern recognition, 2018.

[29] Dawn Song, Kevin Eykholt, Ivan Evtimov, Earlence Fernandes, Bo Li, Amir Rahmati, Florian Tramer, Atul Prakash, and Tadayoshi Kohno. Physical adversarial examples for object detectors. In Proc, of the 12th USENIX workshop on offensive technologies (WOOT 18), 2018.

[30] Zelun Kong, Junfeng Guo, Ang Li, and Cong Liu. Physgan: Generating physical-world-resilient adversarial examples for autonomous driving. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 14254–14263, 2020.

[31] Yago Romano Martinez, Brady Carter, Abhijeet Solanki, Wesam Al Amiri, Syed Rafay Hasan, and Terry N Guo. Mitigation of camouflaged adversarial attacks in autonomous vehicles–a case study using carla simulator. arXiv preprint arXiv:2502.05208, 2025.

[32] Wei Wang, Yao Yao, Xin Liu, Xiang Li, Pei Hao, and Ting Zhu. I can see the light: Attacks on autonomous vehicles using invisible lights. In In Proc. of the ACM SIGSAC Conference on Computer and Communications Security, 2021.

[33] Giulio Lovisotto, Henry Turner, Ivo Sluganovic, Martin Strohmeier, and Ivan Martinovic. SLAP: Improving physical adversarial examples with Short-Lived adversarial perturbations. In Proc. of 30th USENIX Security Symposium (USENIX Security 21), 2021.

[34] Chen Yan, Zhijian Xu, Zhanyuan Yin, Stefan Mangard, Xiaoyu Ji, Wenyuan Xu, Kaifa Zhao, Yajin Zhou, Ting Wang, Guofei Gu, et al. Rolling colors: Adversarial laser exploits against traffic light recognition. In 31st USENIX Security Symposium (USENIX Security 22), pages 1957–1974, 2022.

[35] Ben Nassi, Yisroel Mirsky, Dudi Nassi, Raz Ben-Netanel, Oleg Drokin, and Yuval Elovici. Phantom of the ADAS: Securing advanced driver-assistance systems from split-second phantom attacks. In Proc. of the ACM SIGSAC Conference on Computer and Communications Security, CCS '20, New York, NY, USA, 2020. Association for Computing Machinery.

[36] Ce Zhou, Qiben Yan, Yan Shi, and Lichao Sun. DoubleStar: Long-Range attack towards depth estimation based obstacle avoidance in autonomous systems. In Proc. of 31st USENIX Security Symposium (USENIX Security 22), Boston, MA, August 2022. USENIX Association.

[37] Drew Davidson, Hao Wu, Rob Jellinek, Vikas Singh, and Thomas Ristenpart. Controlling UAVs with sensor input spoofing attacks. In Proc. of the 10th USENIX Workshop on Offensive Technologies (WOOT 16), Austin, TX, August 2016. USENIX Association.

[38] Ranjie Duan, Xiaofeng Mao, A. K. Qin, Yuefeng Chen, Shaokai Ye, Yuan He, and Yun Yang. Adversarial laser beam: Effective physical-world attack to dnns in a blink. In Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 16057–16066, 2021.

[39] Abhijeet Solanki, Syed Rafay Hasan, and Terry N Guo. Investigate the effects of laser attack on intelligence of the AV perception. In Proc. of IEEE Computer Society Annual Symposium on VLSI (ISVLSI), 2024.

[40] Timothy Trippel, Ofir Weisse, Wenyuan Xu, Peter Honeyman, and Kevin Fu. Walnut: Waging doubt on the integrity of mems accelerometers with acoustic injection attacks. In Proc. of IEEE European symposium on security and privacy (EuroS&P), 2017.

[41] Shoei Nashimoto, Daisuke Suzuki, Takeshi Sugawara, and Kazuo Sakiyama. Sensor con-fusion: Defeating kalman filter in signal injection attack. In Proc. of Asia Conference on Computer and Communications Security, 2018.

[42] Yunmok Son, Hocheol Shin, Dongkwan Kim, Youngseok Park, Juhwan Noh, Kibum Choi, Jungwoo Choi, and Yongdae Kim. Rocking drones with intentional sound noise on gyroscopic sensors. In Proc. of the 24th USENIX security symposium (USENIX Security 15), 2015.

[43] Xiaoyu Ji, Yushi Cheng, Yuepeng Zhang, Kai Wang, Chen Yan, Wenyuan Xu, and Kevin Fu. Poltergeist: Acoustic adversarial machine learning against cameras and computer vision. In Proc. of the IEEE Symposium on Security and Privacy (SP), 2021.

[44] Yazhou Tu, Zhiqiang Lin, Insup Lee, and Xiali Hei. Injected and delivered: Fabricating implicit control over actuation systems by spoofing inertial sensors. In Proc. of the 27th USENIX security symposium (USENIX Security 18), 2018.

[45] Chen Yan, Guoming Zhang, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. The feasibility of injecting inaudible voice commands to voice assistants. IEEE Transactions on Dependable and Secure Computing, 18(3):1108–1124, 2021.

[46] Man Zhou, Zhan Qin, Xiu Lin, Shengshan Hu, Qian Wang, and Kui Ren. Hidden voice commands: Attacks and defenses on the VCS of autonomous driving cars. IEEE Wireless Communications, 26(5):128–133, 2019.

[47] Wenyuan Xu, Chen Yan, Weibin Jia, Xiaoyu Ji, and Jianhao Liu. Analyzing and enhancing the security of ultrasonic sensors for autonomous vehicles. IEEE Internet of Things Journal, 5(6):5015–5029, 2018.

[48] Yi Zhu, Chenglin Miao, Hongfei Xue, Zhengxiong Li, Yunnan Yu, Wenyao Xu, Lu Su, and Chunming Qiao. Tilemask: A passive-reflection-based attack against mmwave radar object detection in autonomous driving. In Proc. of the ACM SIGSAC Conference on Computer and Communications Security, 2023.

[49] Zhi Sun, Sarankumar Balakrishnan, Lu Su, Arupjyoti Bhuyan, Pu Wang, and Chunming Qiao. Who is in control? practical physical layer attack and defense for mmwave-based sensing in autonomous vehicles. IEEE Transactions on Information Forensics and Security, 16:3199–3214, 2021.

[50] R Spencer Hallyburton, Yupei Liu, Yulong Cao, Z Morley Mao, and Miroslav Pajic. Security analysis of Camera-Lidar fusion against Black-Box attacks on autonomous vehicles. In Proc. of the 31st USENIX Security Symposium (USENIX Security 22), 2022.

[51] Junjie Shen, Jun Yeon Won, Zeyuan Chen, and Qi Alfred Chen. Drift with devil: Security of multi-sensor fusion based localization in high-level autonomous driving under GPS spoofing. In Proc. of the 29th USENIX security symposium (USENIX Security 20), 2020.

[52] Zhiyuan Cheng, Hongjun Choi, Shiwei Feng, James Chenhao Liang, Guanhong Tao, Dongfang Liu, Michael Zuzak, and Xiangyu Zhang. Fusion is not enough: Single modal attack on fusion models for 3D object detection. In Proc. of The 12th International Conference on Learning Representations (ICLR), 2024.

[53] Jelena Kocić, Nenad Jovičić, and Vujo Drndarević. Sensors and sensor fusion in autonomous vehicles. In Proc. of the 26th Telecommunications Forum (TELFOR), 2018.

[54] S. Filippou, A. Achilleos, S. Z. Zukhraf, C. Laoudias, K. Malialis, M. K. Michael, and G. Ellinas. A machine learning approach for detecting GPS location spoofing attacks in autonomous vehicles. In Proc of IEEE 97th Vehicular Technology Conference (VTC2023-Spring), 2023.

[55] Maliha Shabbir, Mohsin Kamal, Zahid Ullah, and Maqsood Muhammad Khan. Securing autonomous vehicles against GPS spoofing attacks: A deep learning approach. IEEE Access, 11:105513–105526, 2023.

[56] Mohsin Kamal, Arnab Barua, Christian Vitale, Christos Laoudias, and Georgios Ellinas. GPS location spoofing attack detection for enhancing the security of autonomous vehicles. In Proc. of IEEE 94th Vehicular Technology Conference (VTC2021-Fall), pages 1–7, 2021.

[57] Zhen Yang, Jun Ying, Junjie Shen, Yiheng Feng, Qi Alfred Chen, Z. Morley Mao, and Henry X. Liu. Anomaly detection against gps spoofing attacks on connected and autonomous vehicles using learning from demonstration. IEEE Transactions on Intelligent Transportation Systems, 24(9):9462–9475, 2023.

[58] Ahmad Mohammadi, Vahid Hemmati, Reza Ahmari, Frederick Owusu-Ambrose, Mahmoud Nabil Mahmoud, and Abdollah Homaifar. Detection of multiple small biased gps spoofing attacks on autonomous vehicles. In Proc. of IEEE 4th International Conference on AI in Cybersecurity (ICAIC), 2025.

[59] Emre İşleyen and Şerif Bahtiyar. GPS spoofing detection on autonomous vehicles with XGBoost. In Proc. of the 9th International Conference on Computer Science and Engineering (UBMK), 2024.

[60] Vishal Dey, Vikramkumar Pudi, Anupam Chattopadhyay, and Yuval Elovici. Security vulnerabilities of unmanned aerial vehicles and countermeasures: An experimental study. In Proc. of IEEE 31st international conference on VLSI design and 17th international conference on embedded systems (VLSID), 2018.

[61] DJI. Dji phantom 4 pro. https://www.dji.com/phantom-4-pro-v2/.

[62] LabSat3. Labsat3 gps simulato. https://www.labsat.co.uk/index.php/en/.

[63] Shiquan Yi, Jiakai Gao, Yang Lyu, Lin Hua, Xinkai Liang, and Quan Pan. A 3D point attacker for lidar-based localization. In Proc. of IEEE 18th International Conference on Control & Automation (ICCA), 2024.

[64] Yiming Li, Congcong Wen, Felix Juefei-Xu, and Chen Feng. Fooling lidar perception via adversarial trajectory perturbation. In Proc. of the IEEE/CVF International Conference on Computer Vision, pages 7898–7907, 2021.

[65] Han Liu, Yuhao Wu, Zhiyuan Yu, Yevgeniy Vorobeychik, and Ning Zhang. Slowlidar: Increasing the latency of lidar-based detection using adversarial examples. In Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.

[66] Alon Zolfi, Moshe Kravchik, Yuval Elovici, and Asaf Shabtai. The translucent patch: A physical and universal attack on object detectors. In Proc. of the IEEE/CVF conference on computer vision and pattern recognition, 2021.

[67] Naman Patel, Prashanth Krishnamurthy, Siddharth Garg, and Farshad Khorrami. Overriding autonomous driving systems using adaptive adversarial billboards. IEEE Transactions on Intelligent Transportation Systems, 23(8):11386–11396, 2021.

[68] Xingshuo Han, Guowen Xu, Yuan Zhou, Xuehuan Yang, Jiwei Li, and Tianwei Zhang. Physical backdoor attacks to lane detection systems in autonomous driving. In Proc. of the 30th ACM International Conference on Multimedia, 2022.

[69] Noriyuki Miura, Tatsuya Machida, Kohei Matsuda, Makoto Nagata, Shoei Nashimoto, and Daisuke Suzuki. A low-cost replica-based

distance-spoofing attack on mmwave FMCW radar. In Proc. of the 3rd ACM Workshop on Attacks and Solutions in Hardware Security Workshop, ASHES'19, New York, NY, USA, 2019.

[70] Rohith Reddy Vennam, Ish Kumar Jain, Kshitiz Bansal, Joshua Orozco, Puja Shukla, Aanjhan Ranganathan, and Dinesh Bharadia. mmspoof: Resilient spoofing of automotive millimeter-wave radars using reflect array. In Proc. of IEEE Symposium on Security and Privacy (SP), 2023.

[71] Yanlong Qiu, Jiaxi Zhang, Tao Sun, Yanjiao Chen, Jin Zhang, and Bo Ji. Waston: Inferring critical information to enable spoofing attacks using cots mmwave radar. IEEE Transactions on Dependable and Secure Computing, 2024.

[72] Zhiyuan Cheng, Hongjun Choi, James Liang, Shiwei Feng, Guanhong Tao, Dongfang Liu, Michael Zuzak, and Xiangyu Zhang. Fusion is not enough: Single modal attacks on fusion models for 3d object detection. arXiv preprint arXiv:2304.14614, 2023.

[73] Zhi Lian, Peng Shi, Cheng-Chew Lim, and Xin Yuan. Fuzzy-model-based lateral control for networked autonomous vehicle systems under hybrid cyber-attacks. IEEE Transactions on Cybernetics, 53(4):2600–2609, 2023.

[74] NYC.gov. Nyc taxi limousine commission trip record data. https://www.nyc.gov/html/tlc/html/about/trip$_{r}ecord_{d}ata.shtml/$.

[75] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. The international journal of robotics research, 32(11):1231–1237, 2013.

[76] Harald Schafer, Eder Santana, Andrew Haden, and Riccardo Biasini. A commute in data: The comma2k19 dataset. arXiv preprint arXiv:1812.05752, 2018.

[77] Sebastian Houben, Johannes Stallkamp, Jan Salmen, Marc Schlipsing, and Christian Igel. Detection of traffic signs in real-world images: The german traffic sign detection benchmark. In Proc. of IEEE international joint conference on neural networks (IJCNN), 2013.

[78] Markus G. Kuhn. An asymmetric security mechanism for navigation signals. In Jessica Fridrich, editor, Information Hiding, pages 239–252, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[79] K. Wesson, M. Rothlisberger, and T. Humphreys. Practical cryptographic civil gps signal authentication. NAVIGATION, 59(3):177–193, 2012.

[80] Kai Jansen, Matthias Schäfer, Daniel Moser, Vincent Lenders, Christina Pöpper, and Jens Schmitt. Crowd-GPS-Sec: Leveraging crowdsourcing to detect and localize GPS spoofing attacks. In Proc. of IEEE Symposium on Security and Privacy (SP), 2018.

[81] Kasper Bonne Rasmussen and Srdjan Čapkun. Realization of RF distance bounding. In Proc. of the 19th USENIX Conference on Security, USENIX Security'10, USA, 2010.

[82] Matthias Schäfer, Vincent Lenders, and Jens Schmitt. Secure track verification. In Proc. of IEEE Symposium on Security and Privacy, 2015.

[83] L. Lazos, Radha Poovendran, and S. Capkun. ROPE: robust position estimation in wireless sensor networks. In Proc. of Fourth International Symposium on Information Processing in Sensor Networks, 2005.

[84] P. Y. Montgomery, T. E. Humphreys, and B. M. Ledvina. Receiver-autonomous spoofing detection: Experimental results of a multi-antenna receiver defense against a portable civil GPS spoofer. In Proc. of the International Technical Meeting of The Institute of Navigation, 2009.

[85] M. L. Psiaki, S. P. Powell, and B. W. O'Hanlon. Gnss spoofing detection using high-frequency antenna motion and carrier-phase data. In Proc. of the 26th international technical meeting of the satellite division of the Institute of Navigation (ION GNSS+ 2013), 2013.

[86] Z. Zhang, M. Trinkle, L. Qian, and H. Li. Quickest detection of GPS spoofing attack. In Proc. of IEEE Military Communications Conference (MILCOM), 2012.

[87] Stefan Brands and David Chaum. Distance-bounding protocols. In Tor Helleseth, editor, Advances in Cryptology — EUROCRYPT '93, pages 344–359, Berlin, Heidelberg, 1994. Springer Berlin Heidelberg.

[88] FLARM. Ads-r and tis-b now in powerflarm. https://www.flarm.com/en/blog/ads-r-and-tis-b-now-in-powerflarm/.

[89] Daniel Moser, Patrick Leu, Vincent Lenders, Aanjhan Ranganathan, Fabio Ricciato, and Srdjan Capkun. Investigation of multi-device location spoofing attacks on air traffic control and possible countermeasures. In Proc. of the 22nd Annual International Conference on Mobile Computing and Networking (MobiCom), New York, NY, USA, 2016.

[90] Lingxuan Hu and David Evans. Using directional antennas to prevent wormhole attacks. In Proc. of the Network and Distributed System Security Symposium (NDSS). The Internet Society, 2004.

[91] Aanjhan Ranganathan, Hildur Ólafsdóttir, and Srdjan Capkun. Spree: A spoofing resistant gps receiver. In Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking, pages 348–360, 2016.

[92] Todd E Humphreys, Jahshan A Bhatti, and Brent M Ledvina. The GPS assimilator: a method for upgrading existing GPS user equipment to improve accuracy, robustness, and resistance to spoofing. In Proc. of the 23rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS), 2010.

[93] Guodong Rong, Byung Hyun Shin, Hadi Tabatabaee, Qiang Lu, Steve Lemke, Mārtiņš Možeiko, Eric Boise, Geehoon Uhm, Mark Gerow, Shalin Mehta, et al. Lgsvl simulator: A high fidelity simulator for autonomous driving. In Proc. of IEEE 23rd International conference on intelligent transportation systems (ITSC), 2020.

[94] Stylianos Filippou, A Achilleos, SZ Zukhraf, Christos Laoudias, Kleanthis Malialis, Maria K Michael, and George Ellinas. A machine learning approach for detecting GPS location spoofing attacks in autonomous vehicles. In Proc. of IEEE 97th Vehicular Technology Conference (VTC2023-Spring), 2023.

[95] Maliha Shabbir, Mohsin Kamal, Zahid Ullah, and Maqsood Muhammad Khan. Securing autonomous vehicles against gps spoofing attacks: A deep learning approach. IEEE Access, 2023.

[96] Sagar Dasgupta, Mizanur Rahman, Mhafuzul Islam, and Mashrur Chowdhury. A sensor fusion-based gnss spoofing attack detection framework for autonomous vehicles. IEEE Transactions on Intelligent Transportation Systems, 23(12):23559–23572, 2022.

[97] Sagar Dasgupta, Kazi Hassan Shakib, and Mizanur Rahman. Experimental validation of sensor fusion-based gnss spoofing attack detection framework for autonomous vehicles. arXiv preprint arXiv:2401.01304, 2024.

[98] Zhen Yang, Jun Ying, Junjie Shen, Yiheng Feng, Qi Alfred Chen, Z Morley Mao, and Henry X Liu. Anomaly detection against gps spoofing attacks on connected and autonomous vehicles using learning from demonstration. IEEE Transactions on Intelligent Transportation Systems, 24(9):9462–9475, 2023.

[99] Sergei Chuprov, Pavel Belyaev, Ruslan Gataullin, Leon Reznik, Evgenii Neverov, and Ilia Viksnin. Robust autonomous vehicle computer-vision-based localization in challenging environmental conditions. Applied Sciences, 13(9):5735, 2023.

[100] Zhongyuan Hau, Soteris Demetriou, and Emil C. Lupu. Using 3D shadows to detect object hiding attacks on autonomous vehicle perception. In Proc. of IEEE Security and Privacy Workshops (SPW), 2022.

[101] Minkyoung Cho, Yulong Cao, Zixiang Zhou, and Z Morley Mao. Adopt: Lidar spoofing attack detection based on point-level temporal consistency. arXiv preprint arXiv:2310.14504, 2023.

[102] Philipp Meyer, Timo Häckel, Teresa Lübeck, Franz Korf, and Thomas C Schmidt. A framework for the systematic assessment of anomaly detectors in time-sensitive automotive networks. arXiv preprint arXiv:2405.01324, 2024.

[103] Saurabh Jha, Shengkun Cui, Subho Banerjee, James Cyriac, Timothy Tsai, Zbigniew Kalbarczyk, and Ravishankar K. Iyer. Ml-driven malware that targets AV safety. In Proc. of 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), 2020.

[104] Mubeena Begum, Gunasekaran Raja, and Mohsen Guizani. Ai-based sensor attack detection and classification for autonomous vehicles in 6G-V2X environment. IEEE Transactions on Vehicular Technology, 2023.

[105] Andreas Finkenzeller, Anshu Mathur, Jan Lauinger, Mohammad Hamad, and Sebastian Steinhorst. Simutack - an attack simulation framework for connected and autonomous vehicles. In Proc. of IEEE 97th Vehicular Technology Conference (VTC2023-Spring), pages 1–7, 2023.

[106] Sri Hrushikesh Varma Bhupathiraju, Jennifer Sheldon, Luke A. Bauer, Vincent Bindschaedler, Takeshi Sugawara, and Sara Rampazzi. EMI-LiDAR: Uncovering vulnerabilities of lidar sensors in autonomous driving setting using electromagnetic interference. In Proc. of the 16th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '23), 2023.

[107] Yulong Cao, Ningfei Wang, Chaowei Xiao, Dawei Yang, Jin Fang, Ruigang Yang, Qi Alfred Chen, Mingyan Liu, and Bo Li. Invisible for both camera and lidar: Security of multi-sensor fusion based perception in autonomous driving under physical-world attacks. In Proc. of IEEE Symposium on Security and Privacy (SP), 2021.

[108] Niels Balemans, Lucas Hooft, Philippe Reiter, Ali Anwar, Jan Steckel, and Siegfried Mercelis. R2L-SLAM: Sensor fusion-driven slam using mmwave radar, lidar and deep neural networks. In 2023 IEEE SENSORS, pages 1–4, 2023.

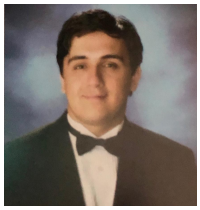
[109] Murad Mehrab Abrar, Raian Islam, Shalaka Satam, Sicong Shao, Salim Hariri, and Pratik Satam. Gps-ids: An anomaly-based gps spoofing attack detection framework for autonomous vehicles. arXiv e-prints, pages arXiv–2405, 2024.
[110] Jianzhi Lou, Qiben Yan, Qing Hui, and Huacheng Zeng. Soundfence: Securing ultrasonic sensors in vehicles using physical-layer defense. In Proc. of 18th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), 2021.
[111] Hamid Ebadi, Mahshid Helali Moghadam, Markus Borg, Gregory Gay, Afonso Fontes, and Kasper Socha. Efficient and effective generation of test cases for pedestrian detection - search-based software testing of baidu apollo in SVL. In Proc. of IEEE International Conference on Artificial Intelligence Testing (AITest), 2021.
[112] Chao Xia, Yanqing Shen, Yuedong Yang, Xiaodong Deng, Shitao Chen, Jingmin Xin, and Nanning Zheng. Onboard sensors-based self-localization for autonomous vehicle with hierarchical map. IEEE Transactions on Cybernetics, 53(7):4218–4231, 2023.
[113] Vysyaraju Manikanta Raju, Vrinda Gupta, and Shailesh Lomate. Performance of open autonomous vehicle platforms: Autoware and apollo. In Proc. of IEEE 5th International Conference for Convergence in Technology (I2CT), 2019.
[114] Gemb Kaljavesi, Tobias Kerbl, Tobias Betz, Kirill Mitkovskii, and Frank Diermeyer. Carla-autoware-bridge: Facilitating autonomous driving research with a unified framework for simulation and module development. arXiv preprint arXiv:2402.11239, 2024.
[115] Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for autonomous driving. In Proc. of the IEEE/CVF conference on computer vision and pattern recognition, 2020.
[116] Baidu Inc. Apollo. http://apollo.auto/, 2017.
[117] Alex H Lang, Sourabh Vora, Holger Caesar, Lubing Zhou, Jiong Yang, and Oscar Beijbom. Pointpillars: Fast encoders for object detection from point clouds. In Proc. of the IEEE/CVF conference on computer vision and pattern recognition, 2019.
[118] Shaoshuai Shi, Xiaogang Wang, and Hongsheng Li. Pointrcnn: 3d object proposal generation and detection from point cloud. In Proc. of the IEEE/CVF conference on computer vision and pattern recognition, 2019.
[119] Mark Harris. Gm cruise snaps up solid-state lidar pioneer strobe inc. https://spectrum.ieee.org/gm-cruise-snaps-up-solidstate-lidar-pioneer-strobe-inc/.
[120] AutoWareAI. Autoware.ai. https://github.com/Autoware-AI/.
[121] Hokuyo. UST-20LX. https://hokuyo-usa.com/products/lidar-obstacle-detection/ust-20lx.
[122] Zizhi Jin, Qinhong Jiang, Xuancun Lu, Chen Yan, Xiaoyu Ji, and Wenyuan Xu. Phantomlidar: Cross-modality signal injection attacks against lidar. arXiv preprint arXiv:2409.17907, 2024.

ABHIJEET SOLANKI (Graduate Student Member, IEEE) received a B.Eng. degree in computer engineering from Tennessee Technological University, Cookeville, TN, USA, in 2020. He is pursuing an M.S. in computer engineering at Tennessee Technological University, where he serves as a Graduate Research Assistant. His primary field of study is cyber-attacks on autonomous vehicle perception systems and mitigation strategies for these threats.

He previously worked as a Senior Software Engineer at Deloitte, Nashville, TN, where he was involved in developing custom software solutions, automating processes, optimizing algorithms, and implementing fraud detection systems for financial clients. His recent publications include "Protecting Electronic Health Records in Transit and at Rest," published in CBMS 2020, and "Investigating the Effects of Laser Attacks on the Intelligence of Autonomous Vehicle Perception," published in ISVLSI 2024. His research interests include perception systems, machine learning, reliable and trustworthy algorithms, and hardware security

WESAM AL AMIRI (Member, IEEE) received his B.S. degree in Communication Engineering from Al-Balqa Applied University, Amman, Jordan, in 2014, and his M.S. and Ph.D. degrees in Electrical and Computer Engineering from Tennessee Technological University, Cookeville, TN, USA, in 2019 and 2024, respectively. He is currently a Post Doc Research Associate and Adjunct Faculty in the Department of Electrical and Computer Engineering, Tennessee Technological University. His research interests include integrated sensing and communication (ISAC), 5G & Beyond, signal processing, cyber-physical systems (CPS) security, and network security.

MARIM MAHMOUD (Student Member, IEEE) received her B.Sc. degree in computer engineering with honors from Qatar University, Doha, Qatar, in 2023. She is now a master's student and Graduate Assistant at Tennessee Technological University in the Department of Electrical and Computer Engineering, focusing on communications applications such as GPS signals and related attack detection and mitigation strategies. Previously, she was a Research Assistant at Qatar University, where she focused on detecting multiple HT-based attacks in IoT-ED without pre-design intervention, enhancing security for modern IoT networks. Her experience also includes roles as an Undergraduate Research Assistant at Qatar University, an intern at Turkish Aerospace working on data communication systems in aviation, and an intern at Qatar Computing Research Institute, where she worked on domain adaptation for natural language processing. Her research interests include signal processing and communications. Ms. Elhanafy served as the vice president of the Qatar University IEEE student branch (2021-2023).



BLAINE SWIEDER received his B.Sc. degree in Mathematics and B.A. in Foreign Languages, with a focus on Spanish, from Tennessee Technological University in 2023. He is currently pursuing his master's degree in computer science with a concentration in Artificial Intelligence from the same university. His research interests lie include scientific computing, autonomous vehicles, and applications of machine learning. He aims to contribute to the advancement of these fields and their real-world applications.

SYED RAFAY HASAN (Senior Member, IEEE) received the B.Eng. degree in electrical engineering from the NED University of Engineering and Technology, Pakistan, and the M.Eng. and Ph.D. degrees in electrical engineering from Concordia University, Montreal, QC, Canada. From 2006 to 2009, he was an Adjunct Faculty Member with Concordia University. From 2009 to 2011, he was a Research Associate with the Ecole Polytechnique de Montreal. Since 2011, he has been with the Electrical and Computer Engineering Department, Tennessee Tech University, Cookeville, TN, USA, where he is currently a full Professor. He has published more than 100 peer-reviewed journal and conference papers. His current research interests include hardware design security in the Internet of Things (IoT), hardware implementation of deep learning, deployment of convolution neural networks in the IoT edge devices, and hardware security issues due to adversarial learning. He received SigmaXi Outstanding Research Award, Faculty Research Award from Tennessee Tech University, the Kinslow Outstanding Research Paper Award from the College of Engineering, Tennessee Tech University, and the Summer Faculty Fellowship Award from the Air force Research Lab (AFRL). Some of his research papers with his graduate students received award including honorable mention in International Symposium on Circuits and Systems (ISCAS). He has received research and teaching funding from NSF, DoE, ICT-funds UAE, AFRL, Qatar National Research Foundation (QNRF) and Intel Inc. He has been part of several funded research projects, as a PI or a Co-PI, worth more than $5.1 million. He has been the Session Chair and Technical Program Committee Member of several IEEE conferences including ISCAS, ICCD, MWSCAS, and NEWCAS, and a regular reviewer for several IEEE Transactions and other journals, including IEEE's TCAD, TCAS-I, TCAS-II, TNNLS, IEEE ACCESS, IEEE Embedded System Letters, Elsevier's Integration, the VLSI Journal, Microelectronics Journal, and IET Circuit Devices and Systems.

TERRY N. GUO (S'96-M'99-SM'10) received the M.S. degree in telecommunications engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 1990 and the Ph.D. degree in communications and electronic systems from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1997. In September 1990, he became a faculty member with UESTC. In January 1997, he joined the Center for Wireless Communications, University of California, San Diego. From December 1999 to January 2002, he was a Research/System Engineer with Golden Bridge Technology, Inc., West Long Branch, NJ, where he was deeply involved in third-generation code-division multiple-access system design, intellectual property development, and standardization activities. From June 2002 to February 2003, he was a Research Engineer with the System Group, Ansoft Corporation, Elmwood Park, NJ, where his major responsibility was software development, with emphasis on functionality modeling of emerging technologies. Since 2004, he has been with the Center for Manufacturing Research, Tennessee Technological University, Cookeville, doing research and development (R&D) and laboratory work. He has more than 15 years of industrial and academic experience in R&D, teaching, and laboratory work. His research interests include wireless communications, statistic signal processing, optimization and its applications, and implementation impact on system performance.