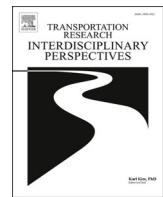




Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives



Vehicle-to-everything (V2X) in the autonomous vehicles domain – A technical review of communication, sensor, and AI technologies for road user safety

Syed Adnan Yusuf^{*}, Arshad Khan, Riad Souissi

Research Department, Elm, Al Raidah Digital City, Riyadh, Saudi Arabia



ARTICLE INFO

Keywords:

Autonomous Vehicles
Vehicle to Everything (V2X)
Millimeter Wave Radars
Vehicles to Infrastructure (V2I)

ABSTRACT

Autonomous vehicles (AV) are rapidly becoming integrated into everyday life, with several countries anticipating their inclusion in public transport networks in the coming years. Safety measures in the context of Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication have been extensively investigated. However, ensuring safety measures for the Vulnerable Road Users (VRUs) such as pedestrians, cyclists, and e-scooter riders remains an area that requires more focused research effort. The existing AV sensor suites offer diverse capabilities, covering blind spots, longer ranges, and resilience to weather conditions, benefiting the V2V and V2I scenarios. Nevertheless, the predominant emphasis has been on communicating and identifying other vehicles, leveraging advanced communication infrastructure for efficient status information exchange. The identification of VRUs introduces several challenges such as localization difficulties, communication limitations, and a lack of network coverage. This review critically assesses the state-of-the-art in the domains of V2X and AV technologies, aiming to enhance the identification, tracking, and localization of VRUs. Additionally, it proposes an end-to-end autonomous vehicle motion control architecture based on a temporal deep learning algorithm. The algorithm incorporates the dynamic behaviors of both visible and non-line-of-sight (NLOS) road users. The work also provides a critical evaluation of various AI technologies to improve the VRU message sharing, identification, tracking and communication domains.

Introduction

The autonomous vehicles (AV) domain is one of the fastest growing areas of the current age due to rapid advances in computing, electronics, sensors and communications technologies. The current developments in computer hardware and sensors have led to significant improvements in the situational awareness of autonomous vehicles (AVs). Additionally, the advancements made in communications, networking, security and performance have facilitated the development of AV platforms that can interact more efficiently with the surrounding infrastructure, as well as other road users. This has contributed to the improvement of the safety and reliability of an average autonomous journey for both the AV and vulnerable road users (VRUs) that include pedestrians and bicyclists. In conventional automotive design and safety use cases, automotive manufacturers extensively consider the VRU safety aspect. The term VRU commonly refers to individuals such as cyclists, pedestrians, and other two-wheeler operators who are more prone to injury or fatality in

vehicle-dominated road spaces. Under the EU ITS Directive, VRUs are defined as “non-motorized road users, such as pedestrians and cyclists as well as motorcyclists and persons with disabilities or reduced mobility and orientation”. With the continued focus on advanced driver assistance systems (ADAS) and other safety features, the safety of VRUs has significantly advanced in conventional vehicles. With the progression of increasing availability of more robust sensors, the preemptive advisory and warning systems for drivers, pedestrians and cyclists have improved substantially in the last decade.

Pedestrians are one of the more common and vulnerable types of various VRUs. According to the Department of Transport, UK ([Road casualty statistics in Great Britain, 2017](#)), 72,993 pedestrian casualties occurred on urban roads in the period of 2017 – 2020. During the past decade, with increasing congestion on roads along with newer modes of transport, the risk to pedestrians by conventional and AVs is likely to increase. Moreover, several studies have reported a steady increase in cyclist-related deaths particularly in heavily built urban areas. However,

* Corresponding author.

E-mail address: syusuf@elm.sa (S. Adnan Yusuf).

studies relate this to an increase in the number of cycle trips in addition to an ever-increasing number of vehicles on the road. The Department of Transport (2020) statistics show a 5 % increase in cyclist fatalities between 2004 and 2020 with a 26 % increase in serious injuries subject to the fact that the pedal cyclist traffic increased by 96 % ([Road casualties in Great Britain: pedal cycle, 2020](#)). Further information on the UK-level statistics shows 73 % of fatalities that lead to serious injuries occurred in urban areas. It is notable that 56 % of fatalities occurred on rural roads. Three-quarters of fatal or serious injury cyclist crashes happened on low-speed roads (30 mph) ([Talbot, et al., 2017](#)). Moreover, with the rapid increase in e-scooter usage, there were 1,280 collisions reported in 2021 compared to 460 in 2020 in the UK ([Road casualties Great Britain: e-Scooter, 2021](#)). These statistics show a significant increase from 2020 when there was just one fatality reported and the serious injuries to pedestrians and cyclists were only 13 and 7, respectively. This shows e-scooters to be a major player in the V2X scope in the presence of AVs in the coming years.

V2X (Vehicle-to-Everything) communication is a technology paradigm that allows the sharing of real-time information between the AV and other road users in the surrounding vicinity. This includes Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), and Vehicle-to-Pedestrian (V2P) with V2P acting as an umbrella term enabling communication with all types of VRUs. V2X describes a vehicle's communication with all other entities in its surrounding environment. This ranges from dynamic entities such as other vehicles, pedestrians, and cyclists, to static elements such as traffic signals, road signs, and road markings. A vehicle communicates with these entities in two distinct ways:

- By receiving feedback from its onboard sensors as a 360-degree area of awareness
- By receiving indirect messages via a communication medium from the infrastructure or other road users

In conventional vehicles, V2X technologies significantly increase a driver's situational awareness, especially in cases where infrastructure occludes other road users. Some well-known V2X systems include cooperative driving/cruise control ([Liu and Kamel, 2016](#)), queue warnings/collision avoidance ([Gelbal, et al., 2017; Vázquez-Gallego, et al., 2019; Ni, et al., 2020; Wu, 2020c](#)), hazard warnings, and extended sensor coverage ([Yamazato, 2015; Lee, et al., 2019; Lee, et al., 2020; Zhou, et al., 2020; Maruta, et al., 2021](#)). For cars driven with V2X, this means improved driver awareness and automotive safety ([Karoui, et al., 2020](#)).

Despite the overarching benefits of V2X technologies to conventionally driven vehicles, with the rapid introduction of autonomous vehicles (AV) on public roads, the role of V2X will transform substantially. As AVs heavily rely on onboard sensors to create a 360-degree area of awareness, they can only view other road users that are predominantly in the direct line of sight. For instance, an AV can be equally unaware of a cyclist who is about to pull in front of it from behind a parked bus than a human driver. Similarly, cross-traffic junctions with viewing obstructions such as hedges reduce the response time of an AV substantially.

Hence, one core objective of this review is:

"To what extent the existing research has explored the development of an integrated communication and perception framework that could increase the situational awareness of other road users especially the VRUs."

In similar scenarios, an integrated V2X is likely to improve the situational awareness of an AV. Despite certain research indicating AV as a relatively low risk medium ([Hulse, et al., 2018](#)), skepticism of the technology has been reported at around 38 % ([Nielsen and Haustein, 2018](#)) compared to the feeling of indifference from a quarter of respondents. A major factor that has so far gained less focus is the public's understanding of the level of risk that pedestrians, cyclists, and other

road users are exposed to when engaging with AV in close-contact urban settings.

This paper presents a detailed review of the existing V2X technologies employed in the AV domain as well as other support infrastructures to improve VRU safety. The scope of this paper will examine the utilization of V2X and Advanced Driver Assistance Systems (ADAS) to improve the safety of AV platforms ([Fig. 1](#)). In doing so, the review will assess the differences in considering AVs in V2X ecosystems by highlighting the technical as well as policy-related differences highlighted in the literature. The work will also consider the role of various sensor technologies that are likely to play a role in improving VRU protection. Moreover, the evolution of the V2X mechanism and the likely route the forthcoming technologies should potential consider. Section II describes various sensors and safety systems onboard vehicles as well as on the VRUs. Section III reports various communication protocols and sensors and reports on ADAS/C-V2X research undertaken so far. Section IV contributes a motion control pipeline in conjunction with various V2X modules. The section describes various AV components that play an essential role in the integration and utilization of the V2X technology along with their strengths and weaknesses. Section V ultimately presents a consolidated explanation of various technologies and research stacks employed leading to a VRU collision avoidance system for AVs. The paper concludes by providing a discussion of potential research venues that could improve the identification and tracking of VRUs in the current state of the art.

V2X safety systems for VRUs

VRU is an umbrella term used for non-motorized road users, such as pedestrians and cyclists, or individuals with disabilities. These users lack an outer shield and hence are at a higher risk of injury on the road in case of a collision. With the rapid advent of newer modes of pedestrian transport, additional VRUs such as electric scooters are now also becoming part of this category. According to the Insurance Institute of Highway Safety (IIHS), pedestrian fatalities count for 17 % of all traffic accident fatalities whereas cyclists account for another 2 % ([Users, 2022](#)). Despite significant advancements in ADAS, including pedestrian detection and safety systems, there still appears to be an increase in pedestrian fatalities over the past 10 years.

The incorporation of safer and robust navigation systems for AVs is becoming more crucial in environments with an increased interaction between AVs and VRUs. Modern AV designs consider VRU mobility patterns and characteristics in a bid to make autonomous navigation safer. One such instance is the incorporation of segregated road-spaces with cyclists since their relatively higher speed requires a quicker response from an AV compared to pedestrians. Moreover, cyclists often show varied behavior patterns such as sudden shifting to pedestrian crossings to avoid crossing main road intersections. Such behaviors make safety incorporation more challenging and generate a larger number of edge cases that an AV's motion planner must consider eliminating collision risks.

According to statistics shared by Brake ([Global road safety statistics | Brake, 2018](#)), more than 1.3 million people die annually from road accidents where 1 in 2 road deaths involve VRUs and 80 % of these belong to developing and rapidly motorizing economies ([Radjou and Kumar, 2018](#)). Over the past few years, planning councils and policymakers have been applying progressive approaches to improve VRU safety by taking measures such as speed limit reductions, allocating dedicated cycling lanes, and smart and signaled crossings and improved visibility measures. These safety measures assist human drivers to get real-time information and guidance cues to be careful in areas with more VRU activity. An average human driver uses their sensory perception to identify developing hazards such as pedestrian trajectories. Conventional V2X technologies improve this perception by understanding a pedestrian's trajectory whereabouts to the vehicle's V2X receiver system. Similarly, cyclist proximity or collision warnings at vehicle blind

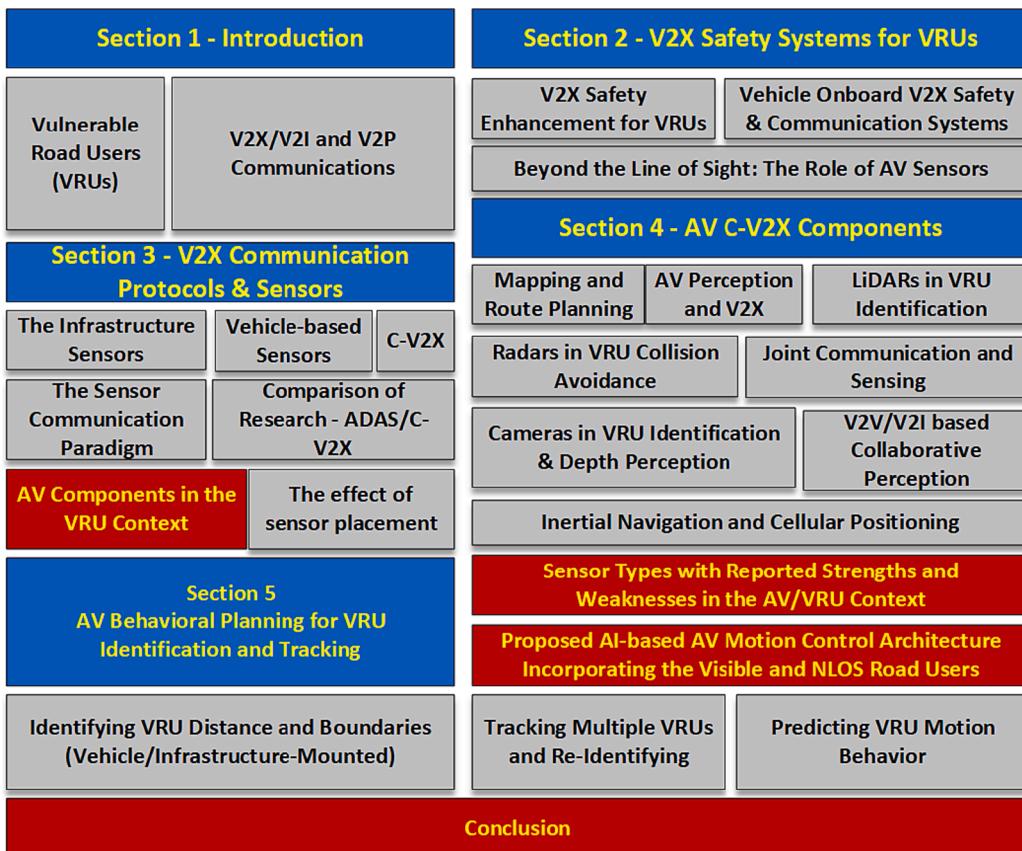


Fig. 1. A diagrammatic organization of this paper along with review sections (blue boxes), sub-topics (gray boxes) and analysis/contributions/proposed architectures (red boxes). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

spots improve reaction times particularly for high-sided vehicles' drivers (Anaya, et al., 2015; Naranjo, et al., 2017). For a human driver, such a collision warning can be a vehicle interface with an alarm or other sensor input. The context changes significantly for AVs in similar circumstances. For example, when there is a blind spot area not covered by an AV's sensors, it is crucial to relay the VRU's location, speed, and other orientation variables to the AV's perception system. This information is essential for estimating the possibility of a collision. The context of this review arises from the fact that other road vehicles and sensors in the vicinity, which have better coverage, can supplement this information. Yet, the state of the art of V2X technologies have so far focused on

- Operational strategies (Pearre and Ribberink, 2019),
- Vehicle platooning applications (Choudhury, 2016; Nardini, 2018; Lekidis and Bouali, 2021; Segata et al., 2021),
- Communication and networking mediums (Pearre and Ribberink, 2019; Abboud et al., 2016; Muhammad and Safdar, 2018; Garcia, et al., 2021),
- Technology testing and validation (Wang, 2019; Jung, 2020; Mafakheri, 2021),
- Commercially available V2X products (Pearre and Ribberink, 2019; Kiela, et al., 2020; Frank and Hawlader, 2021),
- Regulations and standards (Machardy, 2018),
- V2V integration (Zhou, et al., 2020),
- Block chain, security, and privacy aspects (Huang, et al., 2020; Shrestha, et al., 2020).

V2X based safety enhancement for VRUs

Vehicle-to-Pedestrian (V2P) communication represents a specific

category within the broader V2X framework, enabling seamless wireless communication between vehicles and pedestrians. This technology facilitates pedestrians in transmitting their location data effectively to nearby vehicles. Despite sharing an architectural framework akin to V2V and V2I communications, V2P mechanisms exhibit unique characteristics. Notably, the deployment of any V2X node necessitates a connectivity unit at both ends of the communication link. In a typical V2P scenario, an on-vehicle connectivity device establishes a connection with a pedestrian's smartphone, either directly or indirectly. The data exchange process operates through three distinct modes (Malik, 2020).

In the domain of V2X communication, various scenarios shape the level and mode of interaction between vehicles and their surroundings. These scenarios play a pivotal role in determining the effectiveness and coverage of V2X systems. Each of the following three scenarios offers unique approaches and challenges in enhancing communication and safety on the roads.

Scenario 1 – Infrastructure based sensing: In this scenario a roadside network of sensors facilitates V2P communication. Infrastructure based sensing is costlier to set up for most cases and used in more specialized or risky situations such as at signaled junctions, or crossings. This type of setup, however, provides crucial safety information in addition to providing a local consolidation of sensor input from all the actors in the vicinity especially in more challenging situations.

Scenario 2 – Vehicle based sensing: This scenario involves vehicle-based V2X sensors only. It suffers from significantly short coverage such as blind spots, lack of positional information due to GNSS unavailability in dense urban areas and poorer cellular coverage in remote areas. Therefore, in the case of a vehicle not connected to a network (such as C-V2X), even a fully equipped AV is unable to detect objects that are not within its line of sight (NLOS).

Scenario 3 – VRU based sensing: This case includes V2P

communication where each VRU utilizes their smartphone or a body-mounted V2X communication unit to share their position with the surrounding vehicles. This approach is less complex than the first two. Most of the research so far into V2X systems has focused on this type of technology. The effectiveness of this type relies on the quality of the communications network used for message exchange. However, the direct implementation of this genre gets difficult with an increase in the number of users and vehicles in an area.

Vehicle onboard V2X safety and communication systems

The threat posed by a conventional vehicle to a VRU varies based on the vehicle's design and type. For example, sports utility vehicles (SUVs) with their elevated frontal structure are more prone to striking pedestrians, increasing the likelihood of serious injury or fatality. This heightened risk stems from either direct head impacts or the vehicle running over the pedestrian. In conventional vehicles, several safety assessment approaches are reported including trajectory prediction models (Wu, 2019; Zhuang, et al., 2022), V2I-based warning systems (Wang, 2016; Liu et al., 2018) and smart infrastructure integration (Ma, 2019).

There is a growing emphasis on integrating V2X-related safety equipment into modern conventional vehicles. Such vehicles feature onboard sensors in the form of Advanced Driver Assistance Systems (ADAS), serving various functions such as providing information, issuing warnings, or enhancing safety. For example, an in-car Vehicle-to-Pedestrian (V2P) collision avoidance system enables continuous information exchange between pedestrians and vehicles.

A standard V2X system traditionally has the following components (Wang, 2019):

- Vehicle communication device (OBU)
- Pedestrian communication device (smart phone, body-mounted sensor)
- Infrastructure/Roadside Sensing Unit (RSU) (pole/structure mounted sensors)
- Information processing unit (IPU) (edge device or server)

The OBU module facilitates wireless communication between the vehicle and the IPU or the VRUs. There are several modes of communication in a V2X system including Wi-Fi, WiMAX, BTE and cellular networks. Bustos *et al.* presented a computer vision-based pedestrian safety scheme to process street-level images to develop a hazard landscape for VRUs that is updateable at every 15 m (Bustos, et al., 2021). Liu *et al.* presented a radar-based perception method that utilized a 75 GHz radar to detect and track both vehicles and pedestrians and predict collision situations (Liu et al., 2018). A dead reckoning mechanism based on a neural network algorithm to calculate pedestrian's trajectory and predict any impending collision situations has also been reported (Murphy, 2018). In a rapidly developing urban pedestrian scene, real-time discovery of VRUs is of critical importance. One effective approach is the Collective Scanning method, which reduces the time taken by a vehicle to receive responses from pedestrians through various channels. Additionally, it employs an 'Extension Receiving' method to avoid missing Vulnerable Road User (VRU) detections. The method improved the overall detection time by 42.52 ms for the latter technique thereby reducing the overall detection time to 72.13 ms (Fujikami, et al., 2015). He *et al.* reported the latency impact of utilizing WLAN and BTE communication across different V2X communication channels to assess and rectify location inaccuracies in GPS-based pedestrian localization (He et al., 2017). The study's findings indicated that Dedicated Short Range Communication (DSRC), characterized by its low latency, proved to be more effective in scenarios where speed limits exceeded 90 km/h. Furthermore, for speeds up to 108 km/h, GPS accuracy should not deviate by more than 3 m. In recent AV-based collision avoidance scenarios, DSRC-based communication notably demonstrated efficiency,

especially when obstructed line of sight with pedestrians prevented their detection through cameras, radar, or LiDAR. Gelbal *et al.* detected and localized pedestrians in the immediate vicinity of the AV via two DSRC modems (Gelbal et al., 2020). The paper detects and localizes pedestrians when traditional line-of-sight sensors like cameras, radar, and LIDAR fail. The work also introduces a modified version of the elastic band method for real-time collision avoidance and presents model-in-the-loop simulations.

In a generic road traffic scenario, pedestrians and cyclists are the slowest responders to any developing hazards in their vicinity. Hence, a vehicle's awareness of a VRUs accurate position, speed, and orientation in real-time are essential to be integrated with the driver warning systems of conventional vehicles as well as in the motion control logic of AVs. Modern-day smartphones commonly report positioning via the inbuilt GPS combined with several proprietary location calculation methods. However, these methods still do not provide the level of accuracy needed for an AV to pinpoint the position of a VRU. Continuous location monitoring is a power-hungry process due to its continuous background-based pinging. There are various energy-efficient mechanisms reported including a P2P GO method based on the Wi-Fi direct localization. The method reported a 22.4 % improvement in the safe delivery of messages and 36.8 % more energy efficiency compared to the WAVE standard (Lee and Kim, 2016). Substantial work in this area has focused on RSU-IMU-assisted localization (Ma, 2019), LiDAR-to-IMU positioning (Wu, et al., 2020a), GNSS fusion with Inertial Measurement Units (IMUs) and Wheel Speed Sensors via RSUs (Hoang, 2017), integration of inertial sensors with a particle filter to improve vehicular positioning within an Ultra Dense Network (Liu and Liang, 2021).

Beyond the line of sight: The role of AV sensors

Pedestrians' onboard sensor diversity is relatively limited, with smartphones being the most prevalent and often the only devices. Yet, the widespread usage of smartphones, their ubiquitous coverage, and diverse embedded technologies make them the ideal source to localize and track. The literature frequently reports the use of GPS positioning applications on smartphones, serving as collision indicator systems that display warnings and vibration alerts (Hussein, 2016; Liu, et al., 2016; Bampounakis, 2020). To improve GPS signal-related accuracies, several techniques integrate IMU-based corrections to compensate for positioning errors in dense urban areas and tunnels (Hellmers, 2013; Yao, 2017; Wang, et al., 2020a). On the vehicle side, inertial navigation systems with accurate onboard localization identify and broadcast their own location via a communication network. One limitation of such systems is that a large majority utilize smartphone GPS which are power-hungry devices. Li *et al.* have presented a novel power-consumption mechanism that shifts to a coarse-level GPS-based localization mechanism that relies more on the device pedometer (step-counter) and cellular signal strength changes to achieve an average location precision of 92.8 % while conserving 20.8 % energy (Li, et al., 2018).

In recent years, there has been a notable expansion in the deployment and integration of AV platforms. These advancements involve an extensive array of sensors and broader coverage. Thus, V2X safety systems integrated into these platforms have the potential to enhance the safety of VRUs in the surrounding area. For instance, LiDAR or radar sensors onboard an AV offer a considerably wider range and coverage compared to human drivers, enabling more detailed area monitoring. Moreover, the introduction of 5G-based technologies has facilitated real-time communication capabilities between neighboring vehicles. When combined with modern AV sensors, these protocols can substantially improve the identification and tracking of NLOS VRUs. Section III presents a review of V2X communication protocols and sensors. The section further delves into the impact of sensor placement and the related work. Section IV presents a detailed review of AV components in conjunction with how integrated communication and sensing approach contributes

in a NLOS VRU identification and tracking situation.

V2X communication protocols and sensors

Wireless technologies are essential in various safety applications that facilitate interoperable, efficient, and reliable communication in AV-related systems. DSRC is one such technology that originated as an IEEE 802.11p standard to provide Wireless Access for Vehicular Environments (WAVE) for V2X applications. The protocol provided Basic Safety Message (BSM) exchange to share a vehicle's status information that included position, speed, and direction to other users (Yin, et al., 2014). The medium operated on the 75 MHz bands of the 5.9 GHz spectrum with the upper layers and security specified by the IEEE 1609 family. The standard was first approved for use in 2010 and approved in certain 2015 Japanese models before being adopted in some US Cadillac and Volkswagen Golf 8 models in 2017 and 2019, respectively (Boovarahan, 2021). The DSRC had several challenges including channel congestion in high traffic environments, absence of a handshake/ACK in frames broadcast, limited communication due to no internet, inability to receive broadcasted messages and channel leakage (self-interference) (Wu, et al., 2013). The DSRC/WAVE standard utilizes wireless LAN (WLAN) medium to establish DSRC channels to enable vehicle communication on short to medium ranges of up to 300 m. The WAVE standard revolutionized the automotive communication industry as it eliminated the need for an intermediary (server). This not only eliminated latency issues due to direct communication but also extended the applicability of such systems in remote areas with poor cellular

coverage. Soon after the DSRC, another vehicle communication protocol based on cellular networks originated. This protocol was termed C-V2X due to its ability to operate on cellular networks. The goal of this protocol was to improve road safety for more effective communication between vehicles and the infrastructure. This facilitated efficient traffic flow, reducing environmental impact, and providing additional data exchange and passenger information services.

A direct C-V2X allows the transmission of data packets to the receivers while encompassing the following categories (Weber et al., 2019):

- V2V (Vehicle to Vehicle): V2V medium covers safety message exchange between vehicles for collision avoidance, speed control, emergency braking and lane changing.
- V2I (Vehicle to Infrastructure): V2I facilitates the communication between the vehicle and its surrounding infrastructure such as traffic signals, and road signs.
- V2P (Vehicle to Pedestrian): V2P communicates with the nearby VRUs such as pedestrians or cyclists.
- V2N (Vehicle to Network): V2N operates as an infotainment and statistics sharing mechanism where the vehicle can get up-to-date forward traffic or weather updates, or the vehicle sends routine health/service status of the vehicle to the relevant dealership.

Fig. 2 describes a use case covering various V2X components along with sensors to identify and connect various road users.

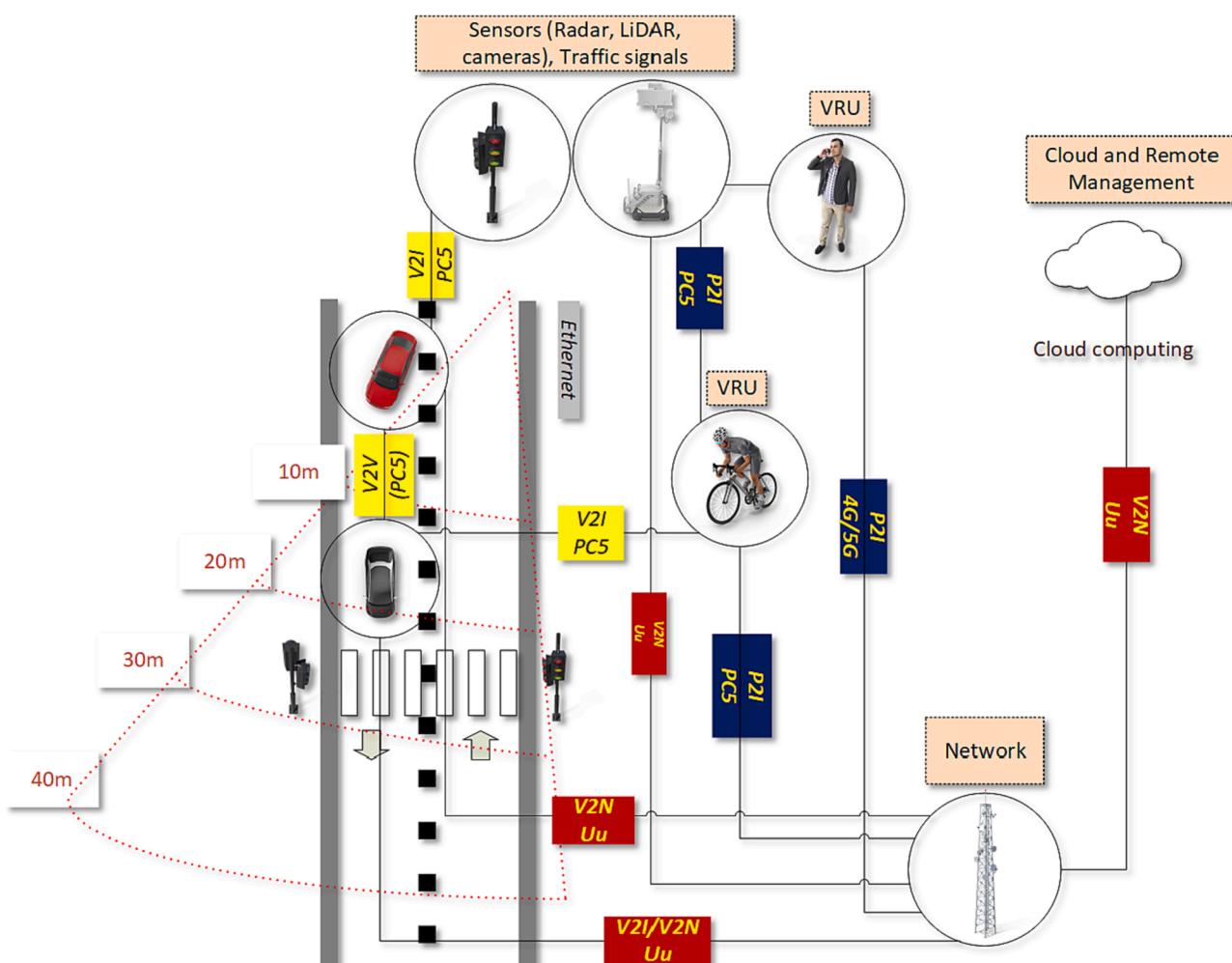


Fig. 2. An indirect C-V2X scenario with the VRU and data exchange in context.

The infrastructure sensors

The infrastructure-based sensor paradigm offers superior coverage due to its flexibility in sensor placement, height, and positioning. This adaptability ensures efficient monitoring of blind spots and detection of VRUs even in NLOS situations. Depending on specific circumstances, the sensors employed can include a variety of technologies such as cameras, radars, and beacon detector systems. Additionally, distance-based sensors such as radars and LiDARs provide precise distance and depth information across different scales. An RSU edge device processes the data collected from these sensors. The on-premises computing of sensor information minimizes the communication overhead of sending all the information to a cloud-based server. Such a localized compute model is termed the Multi-access Edge Computing (MEC) model. The principle of MEC is to localize data processing to local high-performance nodes instead of the cloud. Recent contributions in this domain have reported the utilization of the V2X Cooperative Awareness Messages (CAM) server and VRU-based smartphone GNSS to facilitate real-time status sharing of VRUs with the vehicles (Ru^ß et al., 2016; Napolitano, 2019; Zoghlami et al., 2022). The vehicles share the localization information of the VRUs in real time via these CAM servers.

Vehicle-based sensors

Vehicle-based perception sensors have possibly seen the highest focus of research due to direct application in the ADAS and AV industries. ADAS has gained significant traction in the past 10 years. The provisioning of ADAS is now becoming a norm even in entry-level automotive models. Some of the well-known ADAS technologies include the Automated Emergency Braking Systems (AEBS), adaptive cruise controls, electronic stability control, parking assist, smart navigation, lane assistance, collision avoidance, and artificial intelligence (AI) based assists. In more advanced use cases, AI-based ADAS systems actively report functionalities like automated parking, pedestrian detection, blind spot monitoring, and driver or drowsiness detection (Okuda et al., 2014). ADAS continues to evolve as an increasing number of sensors are integrated into modern vehicles including GPS/GNSS, short-to-long distance radars, cameras, LiDARs, IMUs and SONARs (ultrasonic sensors) (Smith, 2021). However, vehicle-mounted multi-sensor platforms lack the necessary coverage to monitor the relatively small proportions of VRUs adequately due to their limited field of view. A conventional ADAS-equipped vehicle employs a wide array of sensors to create a 360-degree area of awareness. Table 1 presents several ADAS techniques available in modern-day cars alongside the commonly used sensors.

The sensor communication paradigm

Fig. 3 shows an indirect communication mechanism that allows the exchange of information via an on-premises message exchange server. Unlike DSRC, indirect C-V2X allows the cellular network to gather information from several vehicles including non-DSRC/non-cellular vehicles (Ansari, 2018). With the advent of 5G communication, and its better reliability and lower latency, millions of vehicles on cellular networks now incorporate improved situational awareness such as traffic-based speed advisories, GPS map updates, as well as over-the-air (OTA) updates of automotive software. This also facilitated traffic management on larger scales. The original release of this protocol (Release 14/R14) complied with the LTE standard that eventually extended to 5G and 5GNR in Release 15 and 16 respectively. The R16 protocol introduced PC5 and LTE-Uu interfaces to obtain a diversity gain (Lianghai, 2018). The PC5 interface had a shorter range of less than a kilometer and operated in the ITS 5.9 GHz band while providing information on the vehicle's location and speed. The Uu interface provided a longer (>1km) range and provided information from further ahead such as accidents, congestions, etc. The later version (R15) presented

Table 1

A comparative analysis of research and development contributions across ADAS/C-V2X use cases.

Approach (Sensors)	Description	Citations
Cross-traffic assist (GPS/radar)	Sensor fusion to predict collision course with other road users at traffic intersections	(Qi Liu, Liang, et al., 2021)
Intersection movement assist (IMA) (GPS/radar)	Vehicle trajectory estimation for collision estimation	(Miucic, et al., 2018); (Sander et al., 2019); Qi Liu, Liang, et al., 2021)
Emergency Electronic Brake Light (EEBL) warning (Ultrasonics/radar)	Sudden frontal car braking identification	(Miucic, et al., 2018)
Automated Emergency Braking (AEB) (Cameras/radar)	Safety system to identify frontal collisions via trajectories, speed, and orientation estimation	(Sander et al., 2019)
Traffic & route updates (GPS, 5G, radar)	Local area or enroute traffic updates consolidation to issue live reports	(Wu, 2020c)
Blind spot warning (Cameras, radar)	Blind spot driver warning systems against VRUs during lane change, turning or reversing.	(Brandl, 2016); (Jeong et al., 2016); (Naranjo, et al., 2017); (Miucic, et al., 2018); (Maruta, et al., 2021); Qi Liu, Liang, et al., 2021)
Left/right turn assist (Cameras, radar)	Any turn-based assists to minimize collisions from vehicles coming from the opposite direction	(Miucic, et al., 2018); (Sander et al., 2019)
Cooperative lane change (CLC) assists & warning (IMUs, cameras, LiDARs, GNSS, 5G)	A specialized lane-change maneuver based on surrounding traffic density and information to minimize the impact on the overall flow of the traffic	(Brandl, 2016)
VRUs (IMUs, cameras, LiDARs, GNSS, 5G)	VRU's positional and orientation integration with AVs systems to facilitate safer automotive maneuvering	(Hussein, 2016); (Li, et al., 2018); (Malik, 2020); (Parada, 2021)
Adaptive cruise control (ACC) (Cameras, radar, ultrasonic)	Speed adjustable driving to minimize congestion based cellular traffic data	(Liu and Kamel, 2016); (Yadav and Szpytko, 2017)

developments that catered more towards the AV industry and provided network-agnostic, low latency direct communication that could work up to speeds of 500km/hr. Molina-Masegosa and J. Gozalvez presents a comprehensive review of the LTE-V standard based on the PC5 interface (Molina-Masegosa and Gozalvez, 2017). With an improved link budget, this standard performs better than the 802.11p or DSRC alternatives. Most notably, the article presents a detailed analysis of LTE-V mode 4, which does not require any cellular infrastructure support. The mode also supports redundant packet transmission, various sub-channelization schemes, as well as the infrastructure assistance. Recent developments in the Mode 3 domain have introduced enhancements ensuring a more stable resource management, reducing signaling overhead and packet collisions (Aslani et al., 2020); (Sempere-García et al., 2021). As mentioned earlier, Mode 4 needs no support from the cellular infrastructure. This makes the mode more robust, especially in areas with lesser cellular availability. The focus of research in this domain has primarily been on areas such as congestion control (Mansouri et al., 2019), comparison of communication performance, packet transmission, and Modulation and Coding Schemes (Gonzalez-Martín, et al., 2019). Organizations like SAE, IEEE, and ETSI actively support the protocol through security standards (Bazzi, 2019). Although gaining popularity, both the DSRC and C-V2X standards come with their own set of advantages and disadvantages. One aspect of C-V2X facilitated latency-tolerant use cases that included telematics, infotainment, and information safety cases. The 3GPP has recently published its new V2X

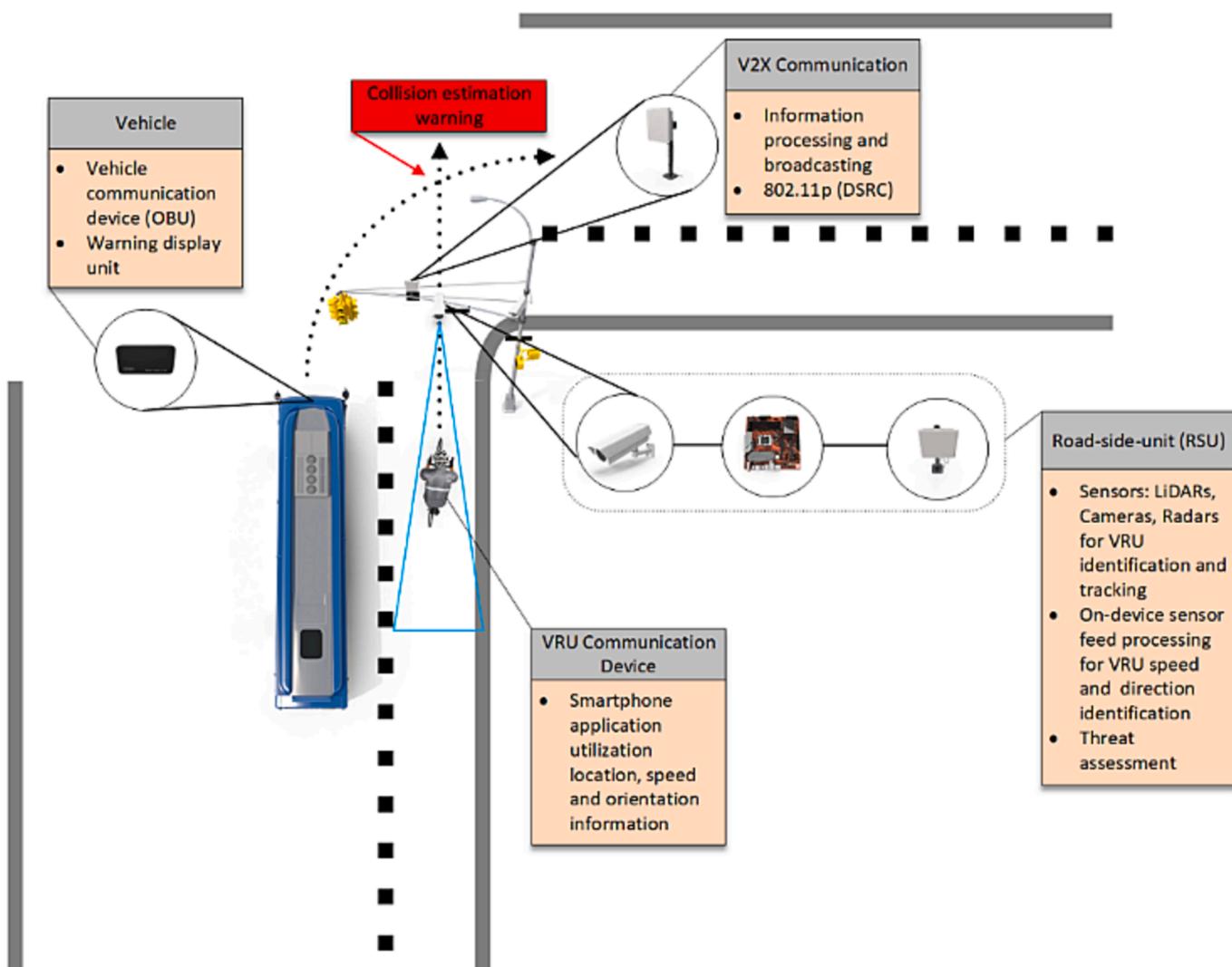


Fig. 3. A use case of V2P protection mechanism utilizing V2I RSU-based sensors and edge devices for cyclist collision detection.

standard known as the 5G New Radio (NR) air interface.

In the work reported in the literature regarding various C-V2X use cases, Sander *et al.* (Sander *et al.*, 2019) re-simulated a 372 left-turn across path/opposite direction (LTAP/OD) accidents. The simulation showed a 59 % crash avoidance if only the turning vehicle was equipped with the Intersection Automated Emergency Braking (AEB), which increased to 77 % in the case of both vehicles equipped with the AEB functionality. Wu *et al.* utilized an RSU to calculate the speed and location of a vehicle pair. They used this information to assess the collision probability of the vehicles and issued intersection collision warnings to the drivers (Wu, *et al.*, 2020b). VRU collision avoidance system based on indirect C-V2X are commonly reported with the VRU identification being carried out via various sensor combinations such as radars and infrastructure-based CCTV cameras. These sensors typically connect to RSUs that are equipped with an image/signal processing edge device via LAN/WLAN or cellular links. The RSU is responsible for the integration of camera and radar-based cyclist identification and distance/depth information from the radar or LiDAR sensors mounted on the signal gantry (Wu, *et al.*, 2020b). Likewise, the RSU receives location and trajectory information from the bus or any vehicle with their planned intersection movement, which in the case shown in Fig. 3 is a right turn. Moreover, depending on the availability, the cyclist may also broadcast their position and speed via a body mounted V2X sensor or a mobile application. Information fusion from the vehicle, infrastructure

and VRU sensor creates an area of awareness. Based on the movement pattern of each of these dynamic parties, the edge device calculates the collision probability of the bus with any VRUs and broadcasts a warning to the bus driver.

The effect of sensor placement

Extending the scope of cooperative V2X perception, the strategic placement of sensors within infrastructure is a critical topic. LiDARs are specifically well-known for their ability to extract rich depth and 3D object information. Hu *et al.* presented a LiDAR sensor placement method aimed at the optimization of the spatial configuration of LiDAR sensors on vehicular platforms. Rather than adhering to the conventional approach of LiDAR installation, data collection, model training, and model performance evaluation, they opted for a unique method. Utilizing 3D bounding boxes, they generated a probabilistic occupancy grid to optimize the placement (Hu, *et al.*, 2021). Dead zones are a common challenge as a LiDAR placed at a specific part of a vehicle's body may create a dead zone on the other side of the vehicle due to the vehicle itself. Kim *et al.* focused more on the dead zones and lower resolution challenges based on sensor placements. Their approach developed a genetic algorithm driven approach where the LiDAR Occupancy Grid (LOG) is encoded as a chromosome. Each LOG receives a fitness score based on the resulting dead zone and resolution of its

placement. This score eliminates the worst performing LOGs and sensor configurations during the evolutionary run (Kim and Park, 2019). Cai et al. developed a LiDAR simulation library in CARLA (Cai, et al., 2022). They used surround, solid state, and prism LiDARs and simulated their digital equivalent in the CARLA simulator and evaluated the placement efficiency simply by the object detection accuracy. For each infrastructure placement, the accuracy was measured via the same V2X/multi-agent perception model. Similar work by Du et al. extended to a range of real-life situations including traffic flow, and speed parameters as well as keeping high-vehicles proportion in the analysis. The team positioned the sensors in both single and dual directional configurations, covering eight scenarios with spacing of 50, 100, 150, and 200 m (Du, et al., 2022). The method used object detection measurements as the evaluation metric and the results showed a significantly low missing probability at 50 m which increased with the increase in the percentage of trucks from 10 % to 75 % with 0.05 to 0.3.

AV components in C-V2X

The design of conventional vehicles based on ADAS systems essentially supports a driver in addition to providing preventative measures such as emergency braking and adaptive cruise control (Table 1). In the case of an AV platform, the context changes since an AV's decision support and motion control principles need a tighter integration of sensors. Extending an AV's sensing capability to include collision possibilities from VRUs brings additional challenges due to the nature of each of these VRU types. In this context, C-V2X represents an information-sharing paradigm model that provides the situational awareness logic for an AV to allow it to incorporate VRU safety. This section presents a critical review of an entire AV pipeline within the V2X context and explains recent research contributions at every relevant stage of the pipeline. An AV platform is equipped with a wide array of sensors that act as the data sources that generate its perception capability. Fig. 4 illustrates the various stages of an AV pipeline.

Mapping and route planning

AV platforms make use of high-definition (HD) maps for navigation and route planning. HD maps may strictly not be needed for autonomous navigation, but they simplify several routing and planning stages. For instance, based on an AV's current position, if a map discovers a traffic cross-section, the V2X module activates a cross-traffic-assist module that will provide crucial spatial-temporal information related to VRUs that are in the vicinity. Another crucial use of HD maps may be in localization where the multi sensor input can be cross validated with the HD map information to get a better position of the platform with reference to the VRUs. In situations where the GPS signal is lost and precise vehicle positioning is crucial, a mapping and route-planning system's ability to rely solely on onboard sensors like cameras proves particularly valuable. Sensor-assisted localization is a commonly explored area for localization. Researchers have recently reported the implementation of Visual Simultaneous Localization and Mapping (SLAM) for autonomous driving and other ADAS-related applications, combining camera-based landmark identification and HD maps (Milz, et al., 2018; Tripathi and Yogamani, 2020; Qin, 2021). A large proportion of these techniques utilize deep learning-based algorithms for the identification and comparison of visual landmarks such as buildings, poles, traffic signs, etc., to localize the vehicle.

These maps create a pre-computed understanding of the world that allows AVs to localize accurately. A standard HD map comprises two core components:

- 3D tiles rendered from depth data obtained from cameras, LiDARs and radars.
- Semantic information to give meaningful labels to various elements in the world.

The latter component elements encompass intersection elements like driving boundaries and virtual lines, as well as traffic signal information elements such as traffic lights and road signs. Additionally, these elements may encompass driving-related information such as crosswalks,

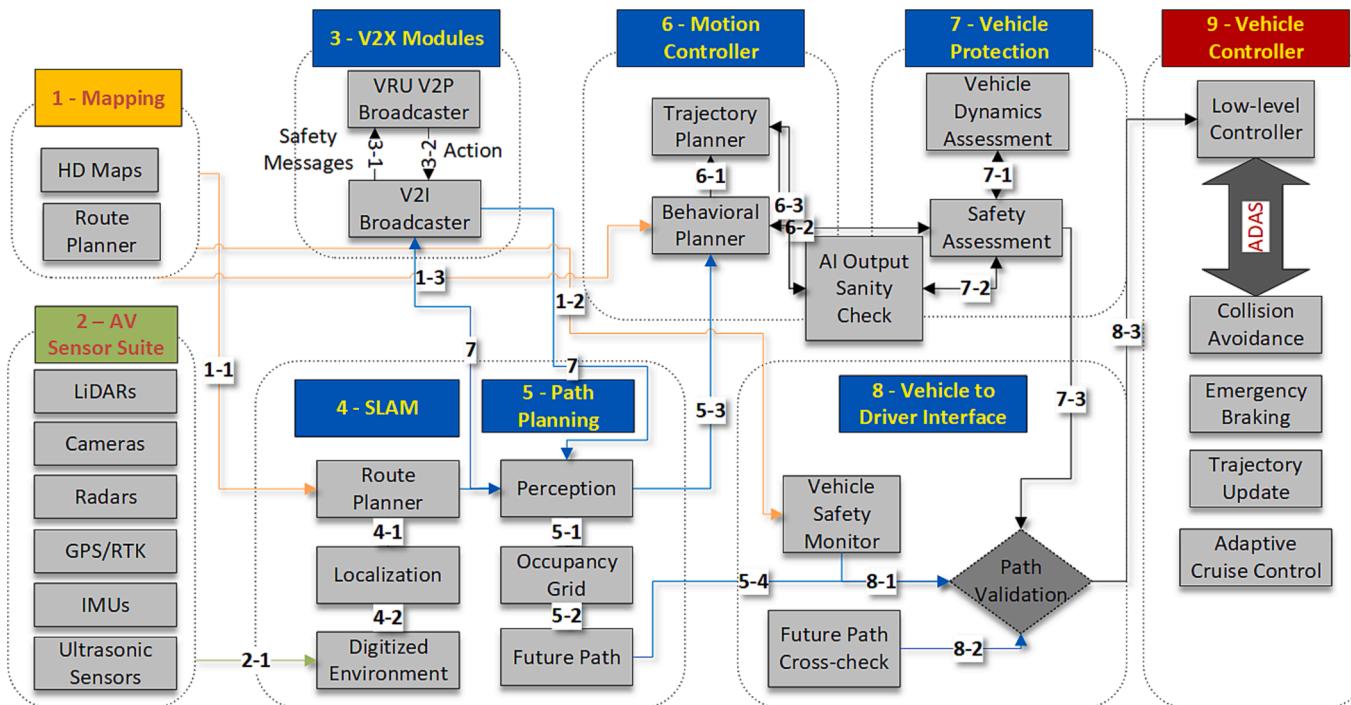


Fig. 4. Comprehensive analysis of various AV components including multi-sensor systems, V2X modules, path/route planning, and vehicle protection strategies in the Vulnerable Road User (VRU) context.

stop lines, buildings, and other static elements. Integrating V2X elements seamlessly is vital to precisely position VRUs and accurately predict their trajectories.

The SLAM module combines the HD map and route information from the mapping module to localize the vehicle and create a digitized environment. The routing information obtained from the mapping module along with the digitized environment performs two core objectives. Firstly, it broadcasts AV's location via the V2I interface to the RSU on the V2X Module. Secondly, the V2X module broadcasts each of the vehicle's location, speed, and orientation-related information to other road users. Moreover, the RSU obtains location information of other VRUs and vehicles to each V2X-bearing vehicle. The SLAM module also shares the localization information with the perception module that is responsible for the generation of a unified understanding of the AV 360-degree surround.

Chou *et al.* have presented the use of ConvNets that use rasterized images to predict pedestrian and cyclist trajectories (Chou, 2020). Carlos Gómez-Huélamo *et al.* have developed a multi-object tracking pipeline based on the semantic information from HD Maps integrated with a Birds Eye View (BEV) Kalman and Hungarian filtering and tracking algorithm. The approach predicts the future trajectories of VRUs based on a Constant Turn Rate and Velocity model (Gómez-Huélamo, *et al.*, 2021). Despite the effectiveness of hybrid sensor-HD map systems for localization, the rudimentary positioning of AV platforms is conventionally done via satellite-based technology commonly termed under the Global Navigation Satellite Systems (GNSS) umbrella. Despite the global effectiveness of the GNSS receivers, several factors affect their standalone integration in an environment where timing, accuracy and availability are of crucial importance to developing a real-time area of awareness for an AV's surroundings.

The AV Perception module consists of various sub-stages including object detection and identification, tracking, traffic and infrastructure recognition, drivable area estimation, depth estimation, and localization. Incorporating VRUs into these sub-stages for safer autonomous vehicle control presents distinct challenges. However, each sub-stage offers crucial information that can enhance the identification and tracking of VRUs. Challenges may vary, from VRU identification and trajectory estimation to the computational demands of each sub-stage.

AV perception and V2X

AVs depend on their perception systems to create an occupancy grid of the surrounding area, encompassing both static and dynamic objects. An AV sensor suite employs an array of sensors to generate the perception of the environment surrounding the vehicle while localizing the vehicle accurately against its rapidly changing scene. For the AV platform to operate safely, real-time processing and understanding of its 360-degree environment are crucial. This section explores the various sensor types used in an AV platform within the V2X/VRU context. Fig. 5 presents a typical sensor integration setup comprising of cameras, LiDARs and radars. With the goal of object detection in mind, prioritizing background subtraction becomes essential as it enables better clustering of 3D point groups from camera depth and LiDAR/Radar point clouds. Moreover, it allows an accurate, pixel-level understanding of each pixel of the scene image. Conventionally, LiDARs are well known for their capability to perform superior background subtraction due to their accurate distance scans (Wang, 2020c; Wen and Jo, 2021). Utilizing long-to-short range radars further enhances the point clustering from the LiDAR feed, improving frontal collision avoidance by removing scanning outliers from both radars and LiDARs (Wang, 2020c). This integration particularly facilitates the identification of smaller-sized objects that radars on their find challenging (Kwon, *et al.*, 2016). Once the foreground is efficiently extracted, the AV is localized based on a combination of GNSS/RTK-based positioning and HD Map referencing commonly known as visual odometry which cross matches critical static objects such as traffic signals, stop lines, etc. via the HD Map reference.

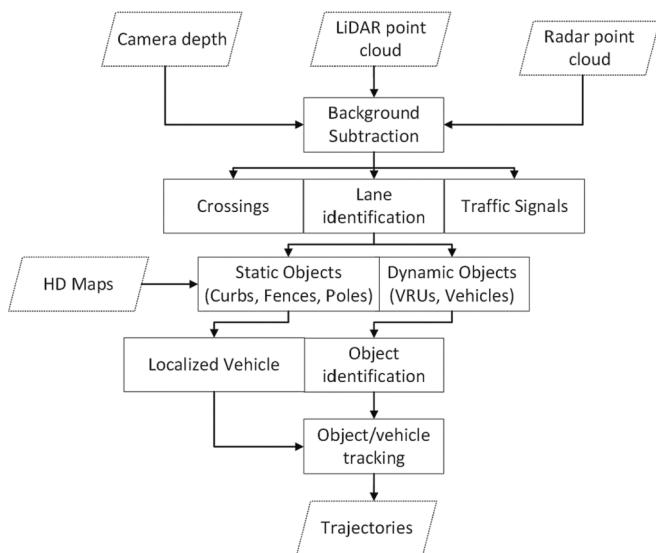


Fig. 5. Sensor integration in an AV perception system for VRU and other dynamic objects' trajectory integration.

The perception system at this stage identifies other dynamic objects such as vehicles, pedestrians, etc. via a combination of local vehicle-based sensors as well as position and orientation feedback from other road users. The consolidation of this information creates a wider behavioral understanding of each of the dynamic objects present in the surrounding area of the AV. This review assesses this stage by examining on-body sensors on various road users, including pedestrians and cyclists, to critically evaluate the current state of the art in VRU safety.

Fig. 6 shows a comparison of strengths and weaknesses of various sensor types in terms of various commonly reported and infrequent edge cases. The commonly reported challenges in an AV's perception system include distance estimation, object speed measurement, lateral resolution, and object classification. The infrequent, edge cases may involve any or more of the cases (e.g., object classification) but in the presence of poor weather, or low visibility conditions. The next sections present a review of various sensor capabilities in terms of VRU identification, coverage/range, distance estimation and speed measurement. A comparison of these pros and cons is further evaluated in Table 2.

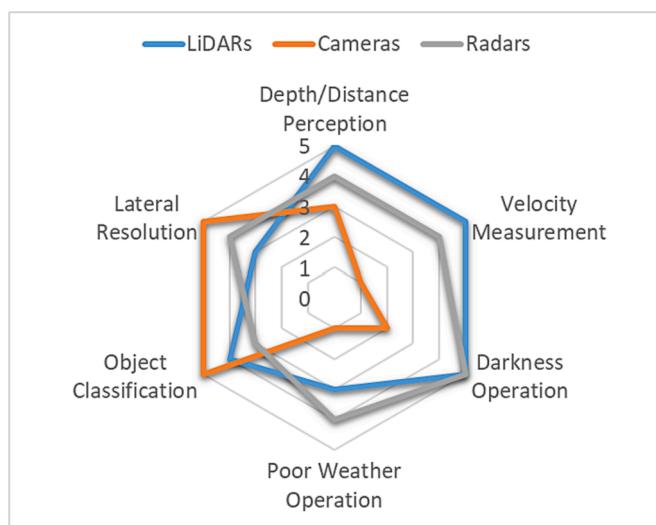


Fig. 6. A comparison on a scale of 0–5 of sensor capabilities in terms of lateral resolution, depth perception, velocity measurement, object classification, weather operability and low-light operation.

Table 2

Strengths and weaknesses of various communication technologies and sensor-based (radar-based) technologies for VRU safety.

Communication Type	Strengths	Weaknesses
6G networks (Wild et al., 2021); (Wymeersch, et al., 2021)	<ul style="list-style-type: none"> Low latency High reliability communication Real-time cooperative safety applications High-density deployments Scalability supported. Network can reach 1 cm detection accuracy (100 + GHz frequency) Camera-free object positioning Object distance/position via LoS Doppler shift for velocity calculation AI/signal processing 	<ul style="list-style-type: none"> Infrastructure equipment deployment Maintenance requirement Cannot identify the object type. Lack of coverage in certain remote areas Increased attenuation at high frequencies Significant bandwidth required. Spectrum resources needed. Data privacy and security considerations
5G Near Radio (Chen, 2021)	<ul style="list-style-type: none"> Scalable/efficient radio interface network and parameters Higher bandwidth of up to 500 MHz Forward compatibility to the next decade 	<ul style="list-style-type: none"> Communication with the network needed for localization and target detection. Designed for communication. RSI is difficult to extract from 5G devices. A joint wave form design for sensing and communication is still missing. Outdoor sensing: More complicated due to more scatter and dynamic objects involved. Existing 5G methods made for shorter ranges. Existing 5G sensing made for single objects. Lack of sensing applications Beamforming: Energy consuming, high running costs Specialized equipment/expertise required. Affected by interference/signal attenuation. Data privacy and security issues
JCAS – single transmitted signal	<ul style="list-style-type: none"> Nearby objects/VRUs detection in real-time (Liu et al., 2018)(Toker and Alsweiss, 2020); (Cui and Dahoun, 2021); (Sheeny, et al., 2021); (Palffy, et al., 2022) All weather/lighting conditions operation (Sheeny, et al., 2021) Distance and speed measurements accuracy (Gelbal et al., 2020) Highest spectral efficiency (Zhang, et al., 2021) No mutual interference Low cost due to shared transmitter and receiver Information sharing via a joint design and optimization 	<ul style="list-style-type: none"> Limited sensing range due to restricted transmission power (Zhang, et al., 2021) Full duplex operation requirement Transmission hardware more complex Interference and clutter are difficult to suppress – unwanted echoes challenging. Clutter suppression techniques from radar to mobile sensing not yet fully developed. Sensing parameter extraction – extraction of sensing parameters challenging due to time, frequency and space variance, random multi-

Table 2 (continued)

Communication Type	Strengths	Weaknesses
	<ul style="list-style-type: none"> architecture. Ability to utilize network resources in the same mobile network (Thomà, 2021) Self-reflected signals from remote radio units used to sense the surrounding environment (Zhang, et al., 2017); (Rahman, 2019) A third receiver exploiting the communication signal for passive sensing. Time difference of arrival and triangulation used for localization. Better target detection capability of distributed MIMO radars (Haimovich et al., 2008) 	<ul style="list-style-type: none"> user access and diversity in resource allocation. (Zhang, et al., 2017) Line-of-sight between the sensor and the object needed. Non-metallic or low-reflective objects may not be detected.

LiDARs in VRU identification

LiDARs are one of the predominant sensors used in the AV industry that generate a depth perception of the surrounding area by emitting lasers. The sensors obtain range information of various objects by receiving laser returns from reflecting object surfaces (Li and Ibanez-Guzman, 2020). Despite being one of the crucial sensors available in an AV design, LiDARs pose certain limitations including inaccuracies on reflective surfaces, data sparsity at longer ranges, degradation at high sun angles or shadows and low operational altitudes (Warren, 2019). Moreover, denser data maps provide better depth information at the expense of the computational complexity involved in object classification tasks. In the VRU context, LiDARs' ability to identify smaller objects due to their shorter wavelength makes them an important sensor. Yet, they suffer from an operating altitude limitation of 500 – 2000 m and their limited usage in nighttime or cloud weather (Widmann, et al., 2000). The sensor is also commonly used to generate HD Maps of areas to assist with AV navigation and motion planning. A large majority of such systems combine various sensors to complement the weaknesses of one with the strengths of the other.

For VRU identification, LiDAR usage has been widely reported in the literature such as in kernel density estimation for pedestrian cluster identification in point clouds that are then translated into location coordinates (Liu et al., 2019). Likewise, Wu et al. propose a multi-LiDAR pedestrian detector as a two-stage pedestrian classifier where each LiDAR presents an SVM-based single class object proposal that is further filtered via an AdaBoost classifier (Wu, et al., 2021). A multi-sensor fusion approach is reported for VRU identification in foggy conditions based on the fusion of information based on a Yolov4 object detector and a LiDAR point cloud fusion (Wu, et al., 2021). Similarly, V2P-based proximity information from pedestrians is relayed and combined with LiDAR-based detections onboard an AV. This hybrid approach eliminates missed VRU detections due to occlusions with the communication system carried by pedestrians broadcasting their location, speed, and direction to the AV's V2P module (Flores, et al., 2019).

LiDARs use rapid, sweeping scans to generate distance-based 3D coordinates known as point clouds. Unlike radars, despite providing crucial distance information even for smaller VRU point clusters, the data for each LiDAR sweep is sparse in nature, irregular and with no order. Hence, the 3D structural information encoded in these scans does not contain as clear information as a conventional camera does. Each point in a LiDAR scan, however, contains features such as reflectivity,

color, normal, etc., which are scale and transformation invariant and hence may prove beneficial for VRU identification. One major drawback of LiDARs is their sensitivity to various weather conditions. In rainy conditions, the water droplets in rain and fog scatter the light emerging from LiDARs thereby reducing the range as well as inducing inaccuracies. Likewise, solid particles in snow and smoke generate false returns leading to unreal obstructions in the sensor's line of sight (Roriz, et al., 2022). Several methods for weather-related LiDAR noise removal are reported in the literature. Some well-known methods in this domain include voxel grid (VG) filters that use outlier removal mechanisms. The VG filters have been reported by several point-cloud improvement methods where point candidates are divided into equally spaced grids that generate one-to-many point mapping between the 3D points and their voxels (Ye, et al., 2020; Quenzel and Behnke, 2021; Wen and Jo, 2021). This method has two sub-methods – hard (Zhou and Tuzel, 2018) and dynamic voxelization (Zhou, et al., 2019; Cui, et al., 2022). The latter method provides more stable detection in point cloud generation by maintaining raw points and voxels. Despite the reported accuracy of voxel grids, they are known to be computationally expensive, especially for complex objects. This makes it further difficult to apply the technique to pedestrian clusters in real-time. Other well-known mechanisms include the statistical outlier removal mechanisms and low-intensity outlier removal techniques (Balta, 2018; Roriz, et al., 2022). Another trend in the V2X genre is that of infrastructure or RSU-based LiDAR feeds. Unlike vehicular LiDARs, RSU-based LiDAR units are generally mounted higher-up and hence are better capable of handling blind spots. Wu et al. have presented an adverse weather conditions handling application based on fixed LiDARs in a roadside use case in windy and snowy weather (Wu, et al., 2020b). The research has presented a ground-filtering method to discard background data in windy conditions based on a ground surface-enhanced point clustering method. Lee et al. have presented LiDAR to LiDAR transition from sunny to adverse weather conditions (Lee, 2022). Rivero et al. used LiDAR as a weather sensor where a 9-month duration of data is collected to identify and model four weather types: clear, rain, fog, and snow. The focus of this research was specifically on the detection of the road asphalt regions (Vargas Rivero, et al., 2020). Roriz et al. proposed an FPGA-assisted weather-related LiDAR signal de-noising method using a Dynamic Light Intensity Outlier Removal algorithm (Roriz, et al., 2022).

On the object classification aspect, the lack of clear spatial traits in LiDAR scans makes it difficult for deep learning models to learn the data against matching class representations. Moreover, the rotation variant nature of LiDAR scans based on which angle of an object was obtained during any specific scan, makes training a deep learning model task even more challenging. Fusion of radar and imaging data with LiDAR scan has frequently been reported to address the deep-learning-related modelling inaccuracies due to data sparsity. A similar approach generates a pixel-level depth feature map (Gao, et al., 2018). The feature map fed into a Convolutional Neural Network (CNN), which learned feature-specific information from this pixel-level data to identify objects in an AV environment in real-time. Despite several attempts to utilize CNN for 3D object classification, the genre has poor handling of sparse datasets due to several reasons. Most notably, CNN's direct application is difficult since unstructured 3D point clouds are irregular (Song, 2020). Moreover, most volumetric 3D CNN methods are complex and hence lead to very large storage and computational costs. Again, this makes it difficult to handle many VRU clusters such as at pedestrian crossings. To resolve these issues, Song et al. presented a CNN-based method that identified 3D objects by transforming the 3D points into Hough Space. The method included a semi-automated point cloud labelling method leading to object 3D Hough space generation and the eventual 3D object classification based on CNN. Alternatively, Kim has used YOLO-based object detection and projected LiDAR feedback on the detections to get a pixel-level classification with 3D object clusters (Kim et al., 2019). Yoshioka et al. presented a Real Ada Boost classification method to identify four object classes in AV LiDAR feed that included cars, pedestrians, cyclists,

and background (Yoshioka, et al., 2018). Capellier et al. focused more on the VRU identification object and focused on a Binary Logistic Classifier to differentiate point clusters representing VRUs from any other objects (Capellier, et al., 2019).

Currently, a large proportion of research into object detection combines various sensors. This technique, however, poses a significant risk where sensor failures may lead to the collapse of the entire perception pipeline. Ideally, each of the sensors should have its own redundancies in place and a global integration layer must not depend on individual sensor feedback (Cui, et al., 2022). Moreover, the possibility of infrastructure mounted LiDARs presents a promising venue for blind spot minimization. Yet, cost, AV standards, long-distance measurements in high-speed journeys, and adverse weather conditions remain some significant challenges (Li and Ibanez-Guzman, 2020). Almost all these challenges can be addressed effectively by the utilization of radar technology in AV.

Radar in VRU collision avoidance

Radar technology in AV operates with a millimeter wave paradigm that provides higher resolution for obstacle detection and centimeter-level accuracy for location and speed determination. The earlier use of radars in the automotive industry was limited to the detection of other road vehicles. The radar sensors used in such scenarios were 2-D capable sensors that only measured the speed and distance of objects. The technology used in imaging radars moved from low channel to high channel count thereby inducing an elevation as well as advanced signal processing and MIMO configurations (Li and Stoica, 2008). Radars are one of the most used sensors for collision avoidance in a V2X architecture. VRUs positions are propagated via a communication link to a roadside terminal and in turn to VRUs as warning signals. A vast majority of work on the communication side of this domain utilizes BLE, Wi-Fi and cellular (4G/5G) signals to communicate with VRU smartphones. Potential arrival areas (PAA) reporting of pedestrian locations in conjunction with an always-on GPS has reported a battery loss of 78 % after a 32-hour test (Li, et al., 2018). Generally, local computation of collision avoidance on user phones is more efficient. On the other hand, performing such high-compute-intensive tasks on smartphones is reported to lead to more battery consumption (Thompson, 2018).

Unlike LiDARs, Radar sensors function effectively in cloudy conditions and during nighttime operations, offering longer ranges. However, due to their longer wavelength, radars can only detect larger objects. This sensor type finds frequent applications in collision avoidance systems. Millimeter-wave radars, commonly operating in a frequency range of 30 – 300 GHz (1 – 10 mm wavelength), are widely used for VRU identification (Liu, 2020). For automotive applications, radars with reported frequencies are generally in the range of 76 – 81 GHz. Millimeter-wave (mm-Wave) radars are frequently utilized in systems for concealed devices detection and posture estimation (Sengupta, 2020), and as a Doppler radar to detect and identify human movement (Chuma and Iano, 2020; Lang, 2020) and high precision human tracking (Cui and Dahnoun, 2021). As per the IHS data, cameras and millimetre-wave/microwave radars form 70 % of automotive collision avoidance sensors. These radars are renowned for their capability to penetrate fog, smoke, and dust, allowing effective operation in challenging environmental conditions (Liu, 2020).

Researchers have extensively discussed pedestrian detection using radars in recent literature with one approach integrating portable GNSS with IMU sensors to establish reference points for automotive radar. Additionally, efforts have been made to integrate LiDAR point clouds with Doppler Effect Regions of Interest (ROIs) from radar scans, addressing the prevalent issue of occlusion (Kwon, et al., 2016; Kwon, 2017) and similar attempts in leveraging depth information from stereo cameras (Palffy et al., 2019). For non-line-of-sight (NLOS) detection, a secondary microwave secondary radar-based detection technique has been proposed (Kawanishi, et al., 2019). Other radar-based pedestrian

detection approaches include Frequency Modulated Continuous Wave (FMCW) used to identify the coherent phase difference of pedestrians with a Doppler FFT (Hyun et al., 2016) elimination of mutual interference (Avdogdu et al., 2018), combining vision-based features with mmWave radars for improved detection (Bi, et al., 2017; Guo, 2018), and incorporation of deep learning techniques to identify pedestrian profiles directly from radar scans (Azizi, 2020); (Palffy, 2020; Wu, et al., 2020b). Despite several attempts on NLOS and partial occlusion challenges, integration of VRUs particularly those hidden completely at NLOS need a robust communication mechanism to improve the collision possibility awareness of AVs.

Radar sensors are ranging sensors like LiDARs that conventionally operate on a 77 GHz frequency with a wavelength of 3.9 mm compared to automotive LiDARs exhibiting smaller wavelengths of 905 – 1550 nm. Due to an even larger wavelength, radar point clouds are sparser than LiDAR point clouds. This characteristic makes it even more difficult to identify two closer or smaller objects. This is generally due to the smaller aperture sizes where most of the signal does not reflect due to specular reflection. Moreover, the lower angular resolution results in object merging which makes it more difficult to cluster two close-by objects (Immoreev and Fedotov, 2002). Despite their weaknesses, radar sensors find widespread automotive applications ranging from cruise controls of ADAS to collision avoidance in emergency braking systems. Traditional radars, also known as 2D radars, typically create an array of readings with each point representing a single object. Imaging radars, alternatively known as 3D radars, offer distinct capabilities.

Conventional automotive radars are somewhat like single-channel LiDARs due to their 2D point cloud return. Despite the somewhat rudimentary and legacy nature of these imaging radars, there has been a substantial amount of research into object identification using imaging radars. Toker and Alsweiss utilized a mmWave radar to distinguish pedestrians from vehicles. This was achieved by detecting distinct limb motions, resulting in varied velocity values, as opposed to the uniform velocity profile generated by rigid vehicular bodies (Toker and Alsweiss, 2020).

On the contrary, the recent 4D genre of radars measures 3D position as well as the velocity of objects. These radars have about a single degree of horizontal and vertical scan resolution. Earlier attempts on radar datasets used these 2D scanning radars lacked Doppler activity that resulted in image-like datasets. Some well-known datasets in this category include Radar RobotCar (Barnes, et al., 2020), RADIATE (Kim, et al., 2020), and MuRan (Kim, et al., 2020). The Frequency-Modulated Continuous-Wave (FMCW) radar datasets in nuScenes and RadarScenes lack elevation information that makes them unsuitable for accurate 3D perception (Caesar, 2020; Schumann, 2021). On the 4D radar, Meyer & Kuschk (Meyer and Kuschk, 2019) Rebut et al. (Rebut, 2022) and Palffy et al., (Palffy, et al., 2022), have reported next-generation radar datasets that include the elevation information to assist with the 3D object detection modelling in the deep learning domain. A large proportion of radar dataset generation activity now focusses on 3D object detection via deep learning algorithms though many report integration with the camera (Meyer and Kuschk, 2019; Nabati and Qi, 2021), [154], and LiDARs (Feng, 2019; Wen and Jo, 2021).

Long-range radars are rapidly finding applications in the autonomous and ADAS domains in use cases such as adaptive cruise control, collision avoidance, emergency braking systems and integrated sensor applications such as improving autonomous vehicles' occupancy grids (Yadav and Szpytko, 2017; Hung, 2019). These radars typically operate on 76 – 77 GHz of the frequency with an angular azimuth and elevation accuracy of up to $\pm 0.1^\circ$ and a range of up to 300 m. Due to their ability to provide four-dimensional measurements including range, Doppler, azimuth and elevation, these radars provide signatures that allow accurate classification of various object types such as cyclists, and pedestrians in complex, fast-speed driving scenarios (Yadav and Szpytko, 2017; Pérez, 2018).

Short-range and surround radars are close-range sensors that provide

blind spot detection (Kuo et al., 2017; Liu, 2017), collision warning (Saleem, 2017; Zakuan, 2017; Arumugam, 2022), parking assist cases (Hung, 2019), and autonomous emergency braking use cases (Flores, et al., 2019). These short-range radars generally provide a wider field-of-view (FOV) of $\pm 90^\circ$ detection and $\pm 75^\circ$ measurement and an operating frequency of between 76 and 81 GHz with a range of up to 180 m for motorcycle detection cases. The surround radar usage includes free-space estimation in AV perception systems, blockage detection, and auto-alignment and interference mitigation. Their elevation detection capability allows usage in curb detection scenarios specifically in lateral collision cases (Flores, et al., 2019). Excluding the short and long-range types, VRU detection and tracking falls in the medium-range category including cross-traffic identification of cyclists, pedestrian detection at crossing intersections and any other on-road VRUs (Arumugam, 2022).

In the medium-range radar category of 4G radars, four signatures range, Doppler, azimuth, and elevation angles, allow better target identification and tracking. Waveform bandwidths, coherent processing interval (CPI) and the aperture of the antenna array determine the Doppler and angular resolutions. For a coherent radar, CPI is termed as the total sampling time. A higher resolution 4D radar imaging requires a longer CPI, greater bandwidth, and aperture size. The Doppler velocity and radar cross sections (RCS) classify road users better. Unlike LiDARs, these radars have lower cost and power consumption along with all-weather operation capabilities that make them ideal for autonomous applications (Li, 2019; Li et al., 2019; Zhou, et al., 2022). To get a more robust and reliable sensor feed for object detection and tracking, the perception stage involves the combination of cameras, radars, and LiDARs with the camera sensors calibrated intrinsically and with LiDAR extrinsically as described by Zhang (Zhang, 2000) and Wang et al. (Wang et al., 2017), respectively.

In a typical AV safety use-case, the circular and medium to short range radars with a range of 40–80 m are commonly reported and address VRU-related use cases including blind spots (Kuo et al., 2017; Liu, 2017; Zakuan, 2017), cross-traffic assist (Brandl, 2016); (Sander et al., 2019), pedestrian detection (Avdogdu et al., 2018; Kawanishi, et al., 2019; Azizi, 2020; Toker and Alsweiss, 2020) etc. (Fig. 7). Long-range radars often combine with cameras and LiDARs to develop more reliable frontal and back collision avoidance and adaptive cruise control systems (Kwon, et al., 2016; Liu and Kamel, 2016; Yadav and Szpytko, 2017; Guo, 2018).

Joint/Integrated communication and sensing (J/I-CAS)

JCAS is a technology paradigm that integrates communication and sensing functionalities in a single entity. This allows the device to perform both communication and sensing together, hence improving reliability, robustness, and security of the system.

For systems below 6 GHz, different communication standards including Wi-Fi, Bluetooth, and cellular 4G/5G technologies with frequencies in the range of 2.4 – 5 GHz including Wi-Fi signals are used. In the sensing context, Wi-Fi signals detect motion and position via the motion-based signals strength variation. In the cellular domain, mobile network operators detect traffic density in specific areas via signal strength. 5G supports ultra-robust, low-latency communication leading to its use in real-time sensing and control applications.

Robots and autonomous vehicles possess accurate geo-localization capabilities along with precise situation recognition abilities. Wi-Fi and Bluetooth signals commonly perform obstacle detection as well as localization and mapping tasks in the AV domain. As discussed earlier, the DSRC technology operates at 5.9 GHz and allows V2V and V2I communication for a wide range of safety applications including collision avoidance, traffic intersection management, and emergency vehicle warning.

Each of the mainstream AV sensors have their own strengths and weaknesses as elaborated previously in Fig. 6. For instance, LiDARs have operation limitations in poor weather conditions leading to a need for

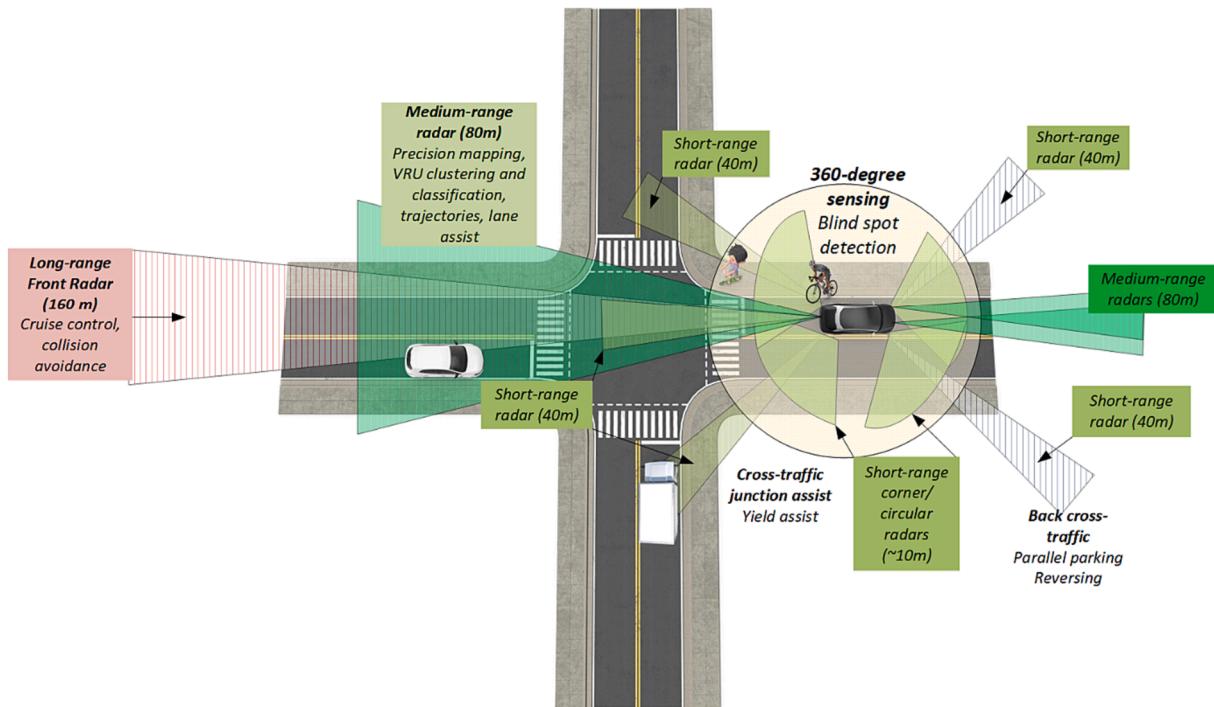


Fig. 7. Coverage and use cases of various AV automotive radars.

radar technology integration. However, the current radars spectrum is not equipped to deal with the large number of AVs expected on roads soon. In addition, mobile spectrum so far cannot cope with the data transfer rates needed for multi-sensor AV platforms introduced an alternative to address this challenge, operating both sensing and communication-based sensing in one spectrum to enable the utilization of that spectrum for both communication and sensing functions. Commonly termed Joint Communication and Sensing (JCAS), the technique exploits the existing communication systems signals for sensing. The technique finds its applications particularly in the AV domain where the vehicle uses the sensors onboard to exchange messages. A capability offered by 6G in this regard is the integration of mobile communication and sensing where the AVs must record their surroundings via the fusion of sensors such as radars, LiDARs, cameras and ultrasonic devices. The cellular spectrum cannot withstand huge bandwidth demands where the positional data from these sensors reach the range of 100 GB/s soon. However, during mobile sensing, significant mobile resources are not needed hence giving rise to the concept of JCAS. A few examples of this mode today are seen in LTE and 5G bands though the device must be connected to the network, send, and receive signals to allow localization. This is more useful in areas with little or no GNSS connectivity, especially in indoor areas. Saur *et al.* demonstrated active localization where wireless 5G interface detected VRUs with up to a meter accuracy along with their trajectories and potential collisions predicted. Another extension of JCAS has been in the Ultra-Wideband (UWB) impulse radio system (IEEE802.15.4z). Telephony, automotive, and electronic manufacturers standardized this system under the Car Connectivity Consortium, and major phone manufacturers have begun utilizing the UWB impulse radio system for localization. This allows enabled devices to localize other objects or mobile devices tagged with beacons. Likewise, the UWB spectrum makes it possible to localize sensor nodes in challenging NLOS use cases such as aircraft cabins (Ninnemann, *et al.*, 2022), exercise activity tracking [169] and indoor individual tracking (Alloulah and Huang, 2019; Shi, *et al.*, 2022). From the traffic perspective of JCAS, Bartoletti *et al.* have explored the use of 5G-V2X side link signals concerning performance bounds on range and speed estimation under different traffic conditions. The research

investigated the use of 5.9 GHz vehicular communication system by exploring the radar sensing ability of side link V2X communication signals (Bartoletti *et al.*, 2022).

Cameras in VRU identification and depth perception

Cameras are possibly the earliest perception sensors used in the automotive industry. The initial focus of imaging devices was on ADAS functionalities such as 360-degree and reverse parking assist, driver awareness, lane departure warning, adaptive cruise control, pedestrian detection, and traffic sign recognition (Dabral, 2014). Like radars, cameras portray a variable set of characteristics when considering a wider scene understanding in the automotive domain. People commonly use these devices for their object classification and depth perception capabilities as standalone sensors or in combination with other sensors (Ming, 2021). However, due to the light-dependent nature of standard imaging devices, common RGB cameras are known to struggle under low visibility, poor weather, and velocity measurement domains.

Researchers have reported the use of multispectral imaging devices for pedestrian detection, integrating non-color information such as depth and thermal signatures to improve detection [175]. On the thermal domain, several attempts have been made using Haar wavelets, AdaBoost, SVM (Qi, 2016), faster R-CNN/CNN (Wagner, *et al.*, 2016; Kim *et al.*, 2018), Random Forests (Lahmyed *et al.*, 2019), and HOG-based feature classification (Rujikietgumjorn and Watcharapinchai, 2017). Most of these techniques originate from a highly compute-intensive image-processing domain where a remotely networked GPU-based processing server results in a performance bottleneck, especially with the increasing proliferation of such image-based V2X nodes. This has recently resulted in the introduction of a more localized compute model termed Multi-access Edge Computing (MEC). The principle of MEC is to localize data processing to local high-performance nodes instead of the cloud. Researchers have proposed the Age-of-Information (AoI) metric for real-time VRU status processing via a cellular network (Emara *et al.*, 2020). Another approach utilizes V2X Cooperative Awareness Messages (CAM) server and VRU-based smartphone GNSS to facilitate real-time status sharing of VRUs with the vehicles (Ruß *et al.*,

2016; Napolitano, 2019); (Zoghlami et al., 2022).

Even though radars and LiDARs are primary distance estimation sensors, obtaining dense depth maps is a computationally expensive task (Ming, 2021). Cameras are one of the most crucial sensors in the AV domain since they offer unmatched object tracking and identification capabilities. Moreover, researchers widely report the capabilities of cameras for rapidly identifying objects in a wide range of scenarios. CNN is one of the most frequently applied deep learning mechanisms with their capabilities in automated feature identification in complex images/scenes and weight-sharing capabilities. Moreover, contrary to standard Multilayer Perceptron (MLP), CNN is shift-invariant with smaller weight sizes and a hierarchical scene understanding architecture. This makes CNN a well-known method in use cases needing rapid object segmentation needs including the AV domain. The technique however is spatially invariant which makes it difficult to learn efficiently from angular variations in 3D objects like those obtained from LiDAR feeds. Yet, the technique has widely been combined with LiDAR point clouds to generate RGB-d based understanding of scenes. There are two main camera types with their own distinct characteristics when it comes to distance measurement, namely monocular and stereovision cameras (Toulminet, 2006). Distance estimation via monocular cameras utilizes correspondence assignment between two image points in two images where the triangulation method estimates the distance of each match. Hence, the accuracy depends upon the match's accuracy and correctness. The monocular method, on the other hand, is a process of estimating the distance of each pixel of an RGB image relative to the camera. Aladem & Rawashdeh developed an encoder-decoder architecture based on CNN to use the single task network to achieve semantic segmentation and depth prediction from a single image (Aladem and Rawashdeh, 2020). Naisen et al. implemented the Shift-RCNN (Faster RCNN) method with 3D object dimension and localized orientation prediction along the camera projection matrix to predict the car, pedestrian, and vehicle classes (Naiden, et al., 2019). Other CNN variants for monocular images-based depth estimation include FastMDE (Dao et al., 2022), and Bayesian cue integration (Mumuni and Mumuni, 2022). Repala and Dubey presented a dual-CNN unsupervised architecture for unsupervised depth estimation in monocular images. The DNM6 and DNM12, 6- and 12-loss models utilized two CNN for each of the left and right stereo-pair images to predict the respective disparities. The results from the DNM12 model showed better performance than the DNM6 (Repala and Dubey, 2019). Li et al. presented a stereo R-CNN to detect and associate objects in right and left images in a simultaneous manner (Li, Chen and Shen, 2019). The method claimed a 30 % improvement on both 3D detection and localization tasks.

V2V/V2I based collaborative perception.

The research community has extensively reported coordinated route planning via V2X. Establishing a datalink where vehicles with different LOS coverage share their scanning features enhances visibility, particularly for VRUs and conventional vehicles. One approach is to pool various permutation invariant operations where Chen et al. presented a Multi-view Object Detection Network (MV3D) integrating the birds-eye-view (BEV) and frontal LiDAR views with the RGB camera images (X. (Chen, 2017)). On the other hand, a view aggregation technique that used multiple, varied-angle object views to generate 3D models by Su et al. was extended by Wang et al. and introduced an innovative concept of sharing neural features via V2X. This extension of the methodology to a different platform enables the creation of a 70-mile information-sharing radius, allowing each vehicle to share and receive information with nearby vehicles (Wang, et al., 2020b). The technique uses a spatially aware graph neural network (GNN) to combine the data received from all the nearby vehicles at various time and viewpoints instances.

Information from surrounding AVs has been categorized into three levels: early fusion, later fusion, and intermediate fusion. Early fusion methods communicate raw sensor data to other vehicles and the

onboard compute capabilities of these vehicles then make predictions based on local processing (Q. (Chen, et al., 2019; Chen, et al., 2019)). Due to the extent of information broadcasted, it is impractical to operate and share this information in real time (Wang, et al., 2020b). Late fusion methods, on the other hand, transmit detection outputs both in temporal and spatial capacity. Rausch et al. proposed a Constant Turn Rate and Acceleration and unscented Kalman Filter predicted the position of the communication partner both in a spatial and temporal domain (Rausch, et al., 2012). Rawashdeh et al. presents a YOLO and Multi-stage DenseNet based bounding box combining and alignment system that identify and localize the vehicles. Furthermore, the study includes vehicle classification, distinguishing between sedan, hatchback, etc., and ultimately calculates the center position of each vehicle (Rawashdeh and Wang, 2018). The performance of the underlying model substantially depends upon each vehicle's performance in the network. Other intermediate approaches have also been reported that share intermediate level features and information with networked vehicles. Chen et al. proposed a voxel-based future fusion method where they saved features in a hash table for non-empty voxels in the 3D LiDAR occupancy map. The approach used Spatial Feature Maps (SFF) that were sparser compared to Voxel Feature Fusion (VFF) and hence more easily compressible with a lesser bandwidth requirement (Chen, et al., 2019; Chen, et al., 2019). Extending on the same concept, OPV2V by Xu et al. put forth a method that utilized minimal bandwidth with high accuracy. The technique compared four LiDAR detectors with all three fusion strategies and compared them with a no-fusion scenario. All four methods generated best results for the intermediate fusion case with VoxelNet improving the AP@IoU accuracy from 0.688 with no fusion to 0.906 for intermediate fusion (R (Xu, et al., 2022; Xu, et al., 2022).

Building upon the intra-vehicular sensor fusion work, we integrated infrastructure elements into the overall picture. Infrastructure-related sensors, being ever-present facilities, were compared to AVs in any setting. This can be specifically useful at key locations such as intersections, and crosswalks. Xu et al. proposed a unified fusion framework termed V2X Vision Transformer. with the infrastructure and vehicle sensors combining intermediate features while encoding and compression the data. On the vehicles side, the V2X-Transformer can then consolidate this information for object detection (Xu, et al., 2022; Xu, et al., 2022). On the dataset aspect, V2V4Real and DAIR-V2X have been reported as two, real-world cooperative perception datasets (Yu, 2022); Runsheng (Xu, et al., 2023). The DAIR-V2X also defines a VIC3D object detection to collaboratively locate and identify based on sensor inputs from both vehicular and infrastructure sensors further demonstrating and improvement of 15 % in AP compared to the single vehicle use case. V2V4Real contributed three benchmarks including 3D detection, tracking and Sim2Real domain adaptation.

Conventional GNSS and augmentation services

Traditional GNSS receivers like those used in standard user smartphones measure the time difference of arrival between a satellite and a mobile device. The signals travel through the ionosphere and atmosphere, causing them to experience a slowdown. Due to errors during this phase, traditional GNSS-only devices normally have accuracies restricted to 2–4 m (Jakowski et al., 2005). Conventional GNSS-based receivers may not be reliable enough for use cases that require pinpoint accuracy and high reliability. Several augmentations to the traditional GNSS have been introduced during the past few years including the satellite-based augmentation systems (SBAS), differential GNSS DGNSS) and Real Time Kinematic (RTK) and Precise Point Positioning (PPP) (Demkowicz, 2022). The conventional GNSS systems calculate positioning by estimating the range between the satellite and an observer. Hence, the apparent range is the observed time of signal transmission between the satellite and the receiver as a multiple of the speed of light. Several types of errors affect the time, including satellite position in orbit, clock errors, satellite biases, receiver hardware

characteristics, and ionospheric and tropospheric effects. (Karaim, 2018). These effects result in inaccuracies of around 3 – 8 m if only relying on satellite signals. GNSS augmentation services play a crucial role in addressing a large proportion of these inaccuracies by providing within-centimeter location accuracy. This section discusses various types of these services in detail.

- **GNSS location service: RTK differential system.**

Real-Time Kinematics (RTK) is a high-performance differential GNSS technique that provides very accurate positioning in the presence of a base station. RTK utilizes carrier-phase and pseudo-range measurements normally recorded at fixed reference locations known as receivers, with the reference network being fixed and the receivers being nonstationary. The fixed base station sends location error corrections to a moving receiver that can be within a 40 km range. The majority of GNSS receivers get their corrections via an internet-based delivery service via the “Networked Transport of RTCM via Internet Protocol” (NTRIP) protocol, satellite, or cellular subscriptions.

The technique employs carrier measurements and a known-location base station to eliminate the primary positioning errors of the rover (GMV, 2011). On its own, Network RTK is known for inaccuracies for vehicular positioning applications including the deterioration in the communication system, lack of coverage and GNSS signal outages (Stephenson, et al., 2012). Walters *et al.* recently studied signal outages within the context of Connected and Autonomous Vehicles (CAV). on the coverage aspect. The work reported only 18 % 4G coverage on UK roads in 2017 whereas in an urban setting the fix loss quality was only 51 % under dense building cover (Walters, 2019). The study found that loss fix quality is not yet sufficient for vehicle tracking even in lightly covered environments when using GNSS RTK-based technologies. The normal accuracy of RTK-based systems remains around 10 cm.

- **Precise Point Positioning**

While the RTK provides corrections for specific locations, Precise Point Positioning (PPP) broadcasts to a larger area with comparatively lower accuracy. This method eliminates or models GNSS-related errors via a single receiver mechanism with corrections transferred via satellites. The solution depends on a network of reference stations that generate GNSS satellite clock and orbit corrections (Bisnath and Gao, 2009). The method combines precise satellite positions and clocks with a dual-frequency GNSS receiver to achieve a centimeter-level accuracy that can be even more accurate in certain post-processing and static modes. A recent study by Du *et al.* evaluated the vulnerabilities of the method; revealing that the most common errors include satellite clock jumps and drifts, bad navigation data uploads, signal power fluctuations, deformed signals, and radio frequency filtration-related errors. (Du, 2021). In the context of AV, the most common limitation of this method is its inability to account for atmospheric calculations in its corrections. While it is globally accessible, even in remote areas, the method may have a longer initialization time, often exceeding 20–30 min, which can make it unsuitable for AV applications (Rizos, et al., 2012).

- **GNSS location based SSR services.**

To mitigate the weaknesses of both RTK and PPP-based systems, one proposed technique is to combine the RTK accuracy and quick initialization times with the larger broadcast technique of PPP. The method requires a base station every 150 km that collects GNSS data to calculate the correction models of both satellites and the atmosphere. A denser reference network is necessary for PPP due to the localized nature of atmospheric corrections. Subscribers receive the corrections through cellular, satellite, or telecommunication services. Generally, a combination of these services performs well to cover a range of use cases from

dense, urban positioning to remote rural areas. The correction model utilizes a message format called Space State Representation (SRR). Recently, SSR/PPP RTK systems have been reported in the context of smartphone positioning, providing Centimeter Level Augmentation Service (CLAS) (Asari et al., 2017; Darugna, et al., 2019; Asari et al., 2020). This approach is expected to bring centimeter-level accuracy to the positioning of VRUs in AV use cases. Table 3 describes the characteristics of various GNSS augmentation services in terms of initialization time, position accuracy, coverage, bandwidth requirement and infrastructure density. Fig. 8 presents a comparison of various communication mediums used to broadcast position error corrections to various user types ranging from high-density urban settings to remote rural areas. Table 4 provides a consolidated comparison of strengths and weaknesses of various sensor types along with potential sensor fusion possibilities.

Inertial navigation and cellular positioning

Despite significant improvements in satellite-driven positioning systems, the domain still presents challenges in the AV domain within enclosed spaces such as tunnels, forest cover or high-density urban areas. This makes any GNSS-based services have poor or no coverage in such areas. Several non-GPS methods are reported in the literature to provide positioning information in areas with poor GNSS reception including local Wi-Fi networks, Inertial Navigation Systems (INS), Cellular Positioning, Bluetooth to sensors and HD maps. Several approaches have been reported in the literature on indoor or GPS-denied positioning use cases including WLAN-based RTK-GPS (Dinh-Van, 2017; Musha and Fujii, 2017), Visible Light Communication (VLC) (Kuo, et al., 2014; Yasir et al., 2014; Xu, 2015), RFID (Bekkali et al., 2007; Ting, 2011; Bai, et al., 2012), Bluetooth (Bekkelien et al., 2012; Jianyong, 2014; Li et al., 2015), Ultra-Wide Band (UWB) (Gigl, 2007; García, 2015), Ultrasonic sensing (Yazici et al., 2011; Yucel, 2012), ZigBee (Zhao, 2008; Liu et al., 2021), IMUs (Hellmers, 2013; Yao, 2017) and computer vision.

VLC uses LED-based lighting sources to provide accurate indoor positioning by modulating data transport via LED sources. The technology is well known for its low power architecture, larger bandwidth support, and secure mode of transport. Despite being very accurate, the technology has its shortcomings including interference issues with other ambient light sources. The technology is reported to have integration challenges with Wi-Fi systems. VLC systems are also known for issues such as atmospheric absorption, shadowing, and beam dispersion. Most importantly, the source and receiver must be within the line of sight which is one of the most prominent shortcomings in the V2X context.

IMUs stand out as one of the most extensively studied non-GNSS sensors, owing to their multifaceted localization and navigation capabilities. These abilities stem from the precise measurements of linear accelerations provided by accelerometers and the rotational movements captured by gyroscopes integrated within IMUs. Liu *et al.* proposed an anti-collision system based on a combination of IMU, RTK, and HD Map-based location calculations for V2P (Qi Liu, Liang, et al., 2021). A large majority of sensor fusion technologies for non-GPS location calculations include positioning calculation based purely on IMUs (Wang, 2016), LiDARs (Eckelmann, 2017); (Wang, et al., 2020a), cameras and HD maps (Toledo-Moreo, 2018); (Xiao, 2019).

Table 3
Comparison of various GNSS augmentation services.

GNSS Augmentation	RTK	RTK-PPP	PPP
Initialization Time (seconds)	1	60	1800
Post initialization accuracy (cm)	~1	2–8	3–10
Coverage	Local	Regional	Global
Bandwidth requirement	High	Moderate	Low
Infrastructure Density (km)	10	100	1000

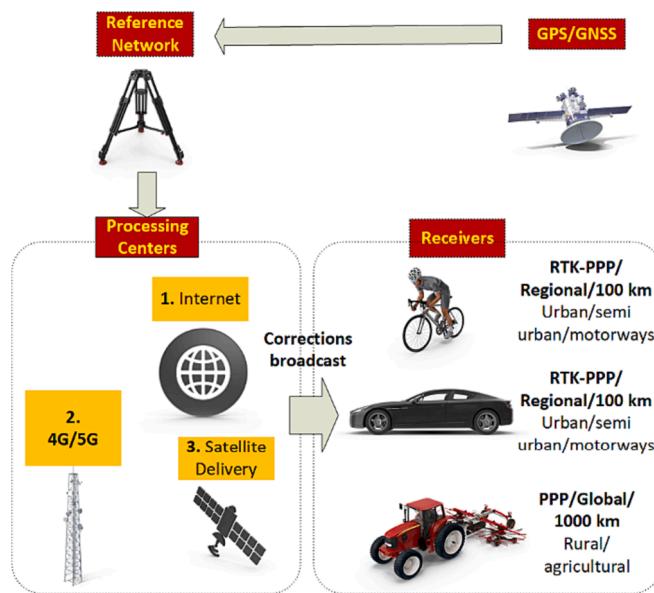


Fig. 8. Comparison of various communication mediums used to broadcast position error corrections.

LTE-V2X was one of the earliest systems that used radio signal-based mechanisms to improve location accuracy (Kong, 2019; Qi Liu, Song, et al., 2021). Moving forward from LTE-V2X, the advent of 5G NR-V2X is now considered a pivotal enabling technology in connected and autonomous mobility sectors. This has increased the location accuracy of the carrier vehicle from > 1 m in LTE-V2X to 0.1 m while providing subcarrier spacing of up to 240 kHz compared to the limited 15 kHz LTE-V2X with increased reliability of 99.9% (Bagheri, et al., 2021). Nr-V2X exploits new positioning methods such as Multicell Round Trip Time (Multi-RTT), Uplink Angle of Arrival (Ul-AoA), and Downlink Angle of Departure (Dl-AoD) to improve the location accuracy (Ghosh, et al., 2019). Qi et al. proposed a User-Equipment (UE) based positioning architecture which contained a Terminal Block based on location information from several sources including satellite, sensor, and cellular network information. The Network Block 5G, RTK and RSU assisted data transmission for the positioning terminal. The third Platform Block provided Real-Time Kinematics (RTK), map database and HD maps-based location compute capability. Finally, the fourth Application Block provided services based on this high-location-accuracy data for various C-V2X-based assisted driving, lane navigation, planning and autonomous navigation tasks.

AV behavioral planning for VRU identification and tracking

In Fig. 9 we have proposed an end-to-end AV pipeline with the integration of the infrastructure and onboard hardware units that combine the localization and motion trajectories of both vehicular and VRU based sensors. The model predicts the future path and motion behavior of the AV based on both the visible road users and NLOS based sensors. The VRUs contribute to a single, unified occupancy grid to generate a consolidated understanding of each of the road users. The occupancy grid draws an object's structural and positional information from both the V2X-connected vehicles and the VRUs. The location and orientation from the vehicular viewpoint generate an area that suffers from information blind spots, particularly for smaller-sized VRUs such as pedestrians. In a standard urban environment, a large proportion of VRUs is likely to be occluded behind fences, poles, or any other road infrastructure creating blind spots. At traffic intersections, RSU-based sensors can minimize a large proportion of such blind spots, but the use of infrastructure-based sensors cannot be made prevalent due to the difficulty and costs involved. To address the shortcoming of both vehicle

and infrastructure-based sensors, VRU-based sensors aim to fill the positional awareness gap of VRUs. As of the current state of the art, provisioning of advanced depth and image-based techniques onboard VRUs does not seem possible due to the disproportionate sizes of these sensors. The smartphone technology however has incorporated many sensors that provide sufficient accuracy in terms of positioning, speed, and orientation of individuals. The most common type of phone-integrated sensors/technologies are GNSS/RTK and IMU such as accelerometers, gyros, and magnetometers. GNSS may still struggle to provide accurate positioning specifically in areas where the sky view is limited; however, IMU-based inertial navigation may still provide methods to compensate for the lack of GNSS-based positioning.

- Identifying VRU distance and boundaries (Vehicle/Infrastructure-mounted)

Semantic segmentation is a deep learning domain that is extensively investigated in computer vision for drivable space and object segmentation. However, the technique on its own suffers substantially from the lack of depth understanding of conventional cameras in cases where, for instance, the pavement matches in color and visual appearance of the road. This shortcoming is often addressed by introducing LiDAR coordinates and combining pixel-level depth by superimposing extrinsically calibrated LiDAR scans with segmented object boundaries (Wang et al., 2017; Wang, 2020c; Wen and Jo, 2021). This also facilitates the identification of smaller objects such as pedestrians and cyclists that are difficult to be identified with standard 4G radar signatures (Kim et al., 2019). The integration of 4G mmWave automotive radars plays a crucial role in collision avoidance at shorter ranges where cross-traffic and on-pavement pedestrians' trajectories can be calculated to estimate collision probabilities. Radars are also frequently reported in cases where even smaller-sized VRUs are identified by combining scan returns with vision information (Lekic and Babic, 2019; Meyer and Kuschk, 2019; Azizi, 2020; Sengupta, 2020; Bai, 2021; Dong, 2021).

Differentiating pedestrians within tight clusters at places such as signaled crossings is one challenge where understanding each, separate road user regardless of occlusion is crucial for tracking. On the vision side, this is addressed by instance segmentation which allows each pedestrian or cyclist to be accurately identified. However, most deep learning instance segmentation methods suffer from the high computational workload and hence are deemed unsuitable for AV/V2X applications processing of several, partially occluded VRUs. Recent reporting in the domain of real-time instance segmentation has been recently reported. Bolya et al. have reported on a two-stage object segmentation method that utilizes a dictionary of non-prototype masks over an entire image and then predicts a set of linear combination coefficients for each instance (pedestrian or cyclists in this case) (Bolya, 2019; Bolya, 2020). Similar single-stage segmentation methods have also been reported for video and image segmentation in domains including car parts segmentation (Cao, 2020; Yusuf et al., 2022).

- Tracking multiple VRUs and re-identifying

Multiple Object Tracking (MOT) is a computer vision technique that analyzes image sequences to establish object motion over connected sequences. After VRU segmentation or detection, tracking is the next stage in the understanding of VRU behavior. Deep-SORT is the most reported method in the tracking domain which was initially reported for vehicle tracking but was then further extended to track other dynamic objects (Hou et al., 2019). The method which is primarily based on a Kalman Filtering and Hungarian algorithm, however, suffers from a significantly high number of false detections and lower accuracy in clustered/occluded objects. Recently, other more efficient tracking mechanisms are reported including Single Shot Multi-Object Tracking (SMOT) and TracTrac (Heyman, 2019; Li, et al., 2020).

Table 4

Role of various sensor types with reported strengths and weaknesses in various combinations.

	Pros	Cons	Potential solution (s)	Distance	Speed	Visibility	Classification & tracking
Radar	Low visibility operation	False detections in clustered groups	The fusing of radar and LiDAR point clouds for more robust VRU cluster identification	(Kuo et al., 2017); (Liu, 2017); (Zakuan, 2017); (Maruta, et al., 2021)	(Hoang, 2017); (Yadav and Szpytko, 2017); (Pérez, 2018)	(Lee, et al., 2017); (Gao, 2021); (Sheeny, et al., 2021)	(Hyun et al., 2016); (Avdogdu et al., 2018); (Kawanishi, et al., 2019); (Azizi, 2020); (Toker and Alsweiss, 2020)
	Resilient to weather conditions	VRU classification inaccuracy due to smaller sizes					
Camera	Accurate object identification	Lack of accurate depth perception	Computer vision-based segmentation, detection and tracking of VRUs.	(Adi and Widodo, 2017); (Salman et al., 2017); (Dandil and Çevik, k.k., 2019); (Zaarane, 2020)	(Wicaksono and Setiyono, 2017); (El Bouziady, 2018); (Tang, 2018)	(Han and Song, 2016); (Alluhaidan and Abdel-Qader, 2018); (Yi et al., 2019)	(Han and Song, 2016); (Wagner, et al., 2016); (Alluhaidan and Abdel-Qader, 2018); (Chebli and Khalifa, 2018); (Tang, 2018); (Yi et al., 2019)
		Poor low visibility operation					
LiDAR	Accurate distance measurement	Difficult object classification	Integration with cameras to improve long-to-medium range VRU recognition.	(Kwon, 2017)	(Dayangac, 2016); (Postica et al., 2016); (Zhang, 2020)	(J. (Wu, 2020c); (Vargas Rivero, et al., 2020); (Sheeny, et al., 2021); (Lee, 2022); (Roriz, et al., 2022)	(Toulimet, 2006); (Yoshioka, et al., 2018); (Feng, 2019); (Ye, et al., 2020); (Quenzel and Behnke, 2021); (Wen and Jo, 2021); (Wu, et al., 2021); (Roriz, et al., 2022)
		Compute-expensive data					
	Better object classification	Sparse dataset	Integration with radars to improve false radar detection and improved collision avoidance				
Radar-Camera-LiDAR Fusion	Accurate identification	Computational overhead	Development of a unified occupancy grid containing person-level identification and tracking	(Dong, 2021); (Long, et al., 2021)	(Ren, 2019); L. (Wang, 2020); (Najman and Zemcik, 2020)	(Lekic and Babic, 2019); Yi, Zhang and Peng, 2019; Z. (Liu, 2021); (Liu and Song, 2021); (Liu, 2021))	(Bi, et al., 2017); (Palffy et al., 2019); (Gao, 2021); (Nabati and Qi, 2021)
	Weather resilience		information shared over the V2X network				
	Long-range object detection						

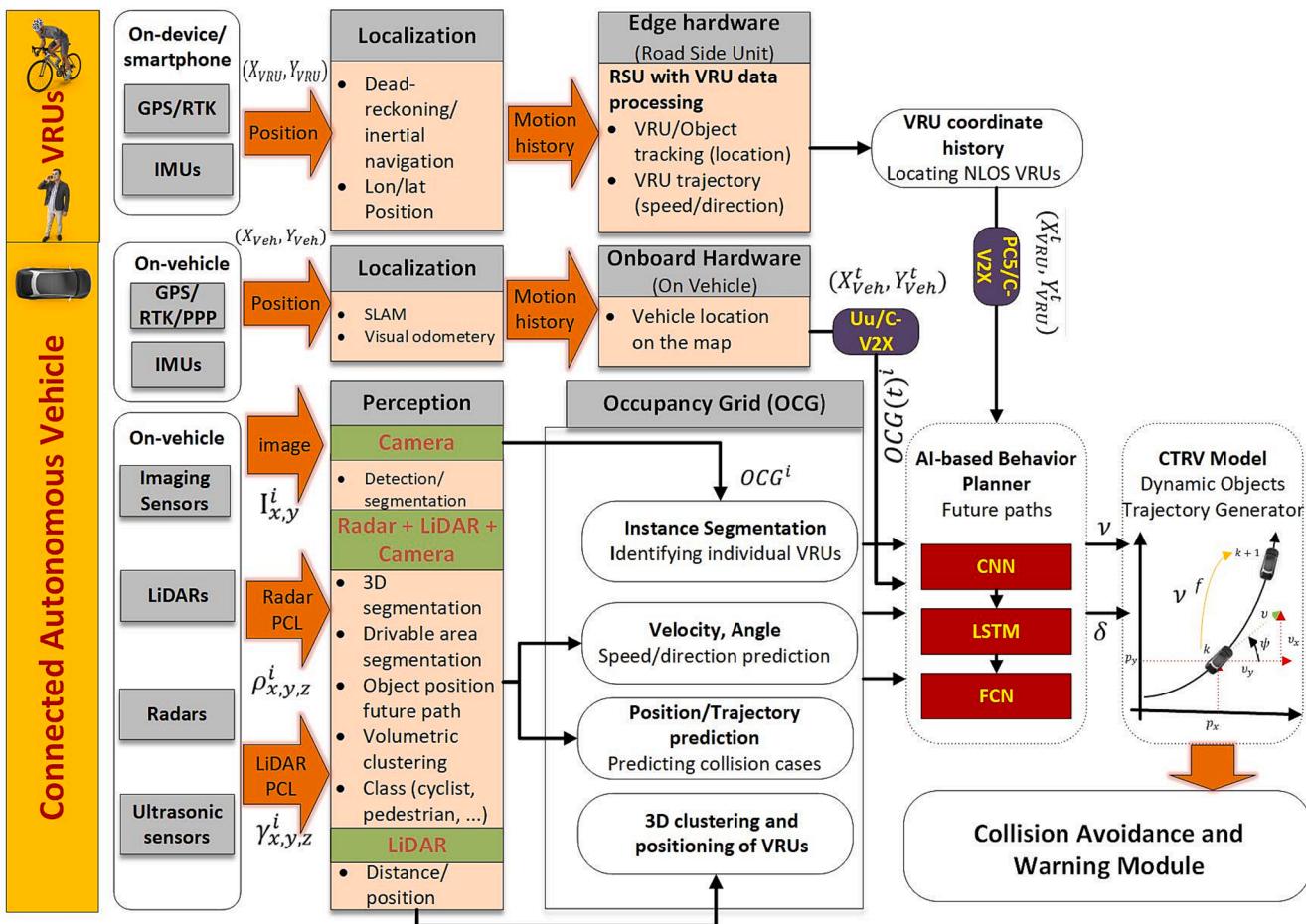


Fig. 9. A proposed consolidation of various technologies/research stacks currently used to develop a threat-assessment and avoidance model for VRUs in the AV domain.

- Predicting VRU motion behavior

Tracking VRUs via vehicle-based sensors can only provide VRU motion information for individuals that are not occluded. NLOS VRUs present the highest level of challenge to an AV as many such may be hidden behind parked cars or other visual obstructions. The sudden appearance of such a VRU may leave less time for the AV perception system to prevent a collision. Inertial navigation technologies on VRU smartphones provide additional location improvement in the absence of reliable GPS reception (Hellmers, 2013); (Yao, 2017). Real-time relay of this information to the vehicles in the vicinity provides the most critical information to prevent collisions. In addition to the position, speed estimation also holds importance in the accurate estimation of the collision parameter of the VRU. Speed estimation of pedestrians is reported under the temporal prediction domain reported by the application of recurrent neural networks (RNN) methodologies such as the Hidden Markov Models (HMM) (Tong, 2019) and Long Short-Term Memory (LSTM) networks (Azizi, 2020). The technique is also well-reported in predicting personnel trajectories in covered areas such as buildings and caves and is known to provide up to centimeter-level accuracy. However, using inbuilt smartphone IMU, the accuracy of dead reckoning for V2P collision avoidance is an area that is yet to be properly investigated. Moreover, due to the vibration-based nature of dead reckoning systems, they are reported to work well for pedestrians. Inertial navigation for cyclists on it's not been commonly reported, and most systems integrate GPS with inertial navigation to improve GPS-only positioning (Shen, 2020).

The ultimate integration of vehicle and infrastructure-based sensor

information consolidates on an occupancy grid which is essentially a cell-based grid array containing the location of each dynamic object. On a temporal scale, it is this 2D/3D representation of AV space that is used to train a machine learning model to facilitate and predict the future path of each of the dynamic objects in the Path Planning stage. The challenge at present is to facilitate the technologies that could make the VRU data available in real time to the AVs in the immediate vicinity. With the availability of a VRU's location and orientation in a vicinity can then be incorporated with the trajectory prediction logic onboard the RSU or the vehicles. A robust behavior prediction algorithm on such an architecture will add an additional protective layer for the VRUs.

Conclusion

The design of AV platforms is rapidly evolving to incorporate advanced sensors and onboard computer hardware with advanced maps and wayfinding capabilities that maximize safety for VRUs. The communication paradigm's efficiency is critical to ensure real-time information sharing among various road users. Future AV development may include sensor costs and power consumption requirements as crucial factors while increasing the range and reducing the form factor. Additionally, advancements in deep learning principles are improving precision and robustness against noisy data, making it possible to identify and track a larger number of dynamic, smaller-sized objects. One direction in this field is the development of advanced sensors that can provide a more comprehensive set of capabilities to act as a safety layer for VRUs.

Another direction is the integration of advanced communication

paradigms that can efficiently share the status of various road users in real-time. This will create an information-sharing network that can maximize the safety of all road users, including VRUs. The development of more advanced and efficient V2X communication protocols and sensors, such as C-V2X, will be essential for enhancing the safety of VRUs. The integration of LiDARs, radars, and cameras for VRU identification and depth perception will also play a vital role in this regard. In this context, the advent of centimeter-level reception using PPP-RTK, and better coverage of cellular networks are also likely to contribute to the GPS/GNSS-based coverage even in denser/shadowed areas.

From the AV sensor point of view, there is an increasingly sophisticated set of algorithms for the behavioral assessment of VRUs. One potential extension to the prediction of VRU trajectories can be investigated towards the integration of temporal deep-learning methodologies such as HMM, LSTM, 3D-CNN or GRUs. 3D CNN is well reported in the literature for their effectiveness in the prediction of vehicular and pedestrian trajectories with applications ranging from video scene understanding to dead-reckoning systems in the IMU domain to track personnel in zero-GPS environments. LSTM due to its ability to temporally contextualize the learning process to remember historic movement patterns may be a potential candidate methodology for situations such as remembering the reappearance of a cyclist being occluded behind a bus to be reappearing on the road again.

As part of this review, we have also proposed an end-to-end AV motion controller architecture that is driven by a temporal deep-neural network that incorporates both visible and NLOS road users. Our ongoing aim is to develop this as a real-world use case and validate it on public roads. To provide concluding remarks on this work, we believe a robust solution for VRU safety can be achieved through a combination of AV technologies, and robust data exchange mechanisms between VRUs via advanced communication technologies, and the use of more reliable positional algorithms.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

- Abboud, K., Omar, H.A. and Zhuang, W. (2016) 'Interworking of DSRC and Cellular Network Technologies for V2X Communications: A Survey', *IEEE Transactions on Vehicular Technology*, 65(12), pp. 9457–9470. Available at: <https://doi.org/10.1109/TVT.2016.2591558>.
- Adi, K., Widodo, C.E., 2017. Distance measurement with a stereo camera. *Int. J. Innov. Res. Adv. Eng.* 4 (11), 24–27.
- Aladem, M., Rawashdeh, S.A., 2020. A single-stream segmentation and depth prediction CNN for autonomous driving. *IEEE Intell. Syst.* 36 (4), 79–85.
- Alloulah, M., Huang, H., 2019. Future millimeter-wave indoor systems: A blueprint for joint communication and sensing. *Computer* 52 (7), 16–24.
- Alluhaidan, M.S., Abdel-Qader, I., 2018. Visibility enhancement in poor weather-tracking of vehicles. In: Proceedings of the International Conference on Scientific Computing (CSC). The Steering Committee of the World Congress in Computer Science, Computer ..., pp. 183–188.
- Anaya, J.J. et al. (2015) 'Vulnerable Road Users Detection Using V2X Communications', in 2015 IEEE 18th International Conference on Intelligent Transportation Systems, pp. 107–112. Available at: <https://doi.org/10.1109/ITSC.2015.26>.
- Ansari, K. (2018) 'Cloud Computing on Cooperative Cars (C4S): An Architecture to Support Navigation-as-a-Service', in 2018 IEEE 11th International Conference on Cloud Computing (CLOUD), pp. 794–801. Available at: <https://doi.org/10.1109/CLOUD.2018.00108>.
- Arumugam, S., et al., 2022. A comprehensive review on automotive antennas for short range radar communications. *Wirel. Pers. Commun.* 1–28.
- Asari, K., Saito, M. and Amitani, H. (2017) 'SSR assist for smartphones with PPP-RTK processing', in Proceedings of the 30th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2017), pp. 130–138.
- Asari, K., Kubo, Y., Sugimoto, S., 2020. Design of GNSS PPP-RTK assistance system and its algorithms for 5G mobile networks. *Transactions of the Institute of Systems, Control and Information Engineers* 33 (1), 31–37.
- Aslani, R., Saberinia, E. and Rasti, M. (2020) 'Resource Allocation for Cellular V2X Networks Mode-3 With Underlay Approach in LTE-V Standard', *IEEE Transactions on Vehicular Technology*, 69(8), pp. 8601–8612. Available at: <https://doi.org/10.1109/TVT.2020.2997853>.
- Avdogdu, C., Garcia, N., Wymeersch, H., 2018. 'Improved pedestrian detection under mutual interference by FMCW radar communications', in 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE 101–105.
- Azizi, M.A.M., et al., 2020. Pedestrian detection using Doppler radar and LSTM neural network. *Int J Artif Intell ISSN* 2252 (8938), 8938.
- Bagheri, H. et al. (2021) '5G NR-V2X: Toward Connected and Cooperative Autonomous Driving', *IEEE Communications Standards Magazine*, 5(1), pp. 48–54. Available at: <https://doi.org/10.1109/MCOMSTD.001.2000069>.
- Bai, J., et al., 2021. Radar transformer: An object classification network based on 4d mmw imaging radar. *Sensors* 21 (11), 3854.
- Bai, Y.B. et al. (2012) 'Overview of RFID-Based Indoor Positioning Technology.', GSR, 2012.
- Balta, H., et al., 2018. Fast statistical outlier removal based method for large 3D point clouds of outdoor environments. *IFAC-PapersOnLine* 51 (22), 348–353.
- Barmpoukakis, S., et al., 2020. Collision avoidance in 5G using MEC and NFV: The vulnerable road user safety use case. *Comput. Netw.* 172, 107150.
- Barnes, D. et al. (2020) 'The Oxford Radar RobotCar Dataset: A Radar Extension to the Oxford RobotCar Dataset', in Proceedings - IEEE International Conference on Robotics and Automation. Institute of Electrical and Electronics Engineers Inc., pp. 6433–6438. Available at: <https://doi.org/10.1109/ICRA40945.2020.9196884>.
- Bartoletti, S., Decarli, N. and Masini, B.M. (2022) 'Sideline 5G-V2X for Integrated Sensing and Communication: the Impact of Resource Allocation', in 2022 IEEE International Conference on Communications Workshops (ICC Workshops), pp. 79–84. Available at: <https://doi.org/10.1109/ICCWorshops53468.2022.9814586>.
- Bazzi, A. (2019) 'Congestion Control Mechanisms in IEEE 802.11p and Sideline C-V2X', in 2019 53rd Asilomar Conference on Signals, Systems, and Computers, pp. 1125–1130. Available at: <https://doi.org/10.1109/IEEECONF44664.2019.9048738>.
- Bekkali, A., Sanson, H. and Matsumoto, M. (2007) 'RFID indoor positioning based on probabilistic RFID map and Kalman filtering', in Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2007). IEEE, p. 21.
- Bekkelien, A., Deriaz, M., Marchand-Maillet, S., 2012. Bluetooth indoor positioning. University of Geneva [Preprint]. Master's thesis.
- Bi, X. et al. (2017) A new method of target detection based on autonomous radar and camera data fusion. SAE Technical Paper.
- Bisnath, S., Gao, Y., 2009. Precise point positioning. *GPS World* 20 (4), 43–50.
- Bolya, D., et al., 2019. Yolact: Real-time instance segmentation. In: In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9157–9166.
- Bolya, D., et al., 2020. Yolact++: Better real-time instance segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [preprint].
- Boavarahan, N., 2021. Vehicle to everything an introduction. *International Journal of Research Publication and Reviews* 2582 (5), 7421.
- Brandl, O., 2016. 'V2X traffic management', *e & i. Elektrotechnik Und Informationstechnik* 133 (7), 353–355.
- Bustos, C. et al. (2021) 'Explainable, automated urban interventions to improve pedestrian and vehicle safety', *Transportation Research Part C: Emerging Technologies*, 125, p. 103018. Available at: <https://doi.org/10.1016/j.trc.2021.103018>.
- Caeser, H., et al., 2020. nuscenes: A multimodal dataset for autonomous driving. In: In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11621–11631.
- Cai, X. et al. (2022) 'Analyzing infrastructure lidar placement with realistic lidar', arxiv.org [Preprint]. Available at: <https://arxiv.org/abs/2211.15975> (Accessed: 27 September 2023).
- Cao, J., et al., 2020. Sipmask: Spatial information preservation for fast image and video instance segmentation. *European Conference on Computer Vision*. Springer 1–18.
- Capellier, E. et al. (2019) 'Evidential deep learning for arbitrary LIDAR object classification in the context of autonomous driving', in IEEE Intelligent Vehicles Symposium, Proceedings. Institute of Electrical and Electronics Engineers Inc., pp. 1304–1311. Available at: <https://doi.org/10.1109/IVS.2019.8813846>.
- Chebli, K., Khalifa, A.B., 2018. Pedestrian detection based on background compensation with block-matching algorithm. In: In 2018 15th International Multi-Conference on Systems, Signals & Devices (SSD). IEEE, pp. 497–501.
- Chen, X., et al., 2017. Multi-view 3d object detection network for autonomous driving. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Available at.
- Chen, Y., et al., 2021. Radio sensing using 5G signals: concepts, state of the art, and challenges. *IEEE Internet Things J.* 9 (2), 1037–1052.
- Chen, Q. et al. (2019) 'Cooper: Cooperative perception for connected autonomous vehicles based on 3d point clouds', in 39th International Conference on Distributed Computing Systems (ICDCS). Available at: <https://ieeexplore.ieee.org/abstract/document/8885377/> (Accessed: 27 September 2023).
- Chen, Q. et al. (2019) 'F-cooper: Feature based cooperative perception for autonomous vehicle edge computing system using 3D point clouds', dl.acm.orgQ Chen, X Ma, S Tang, J Guo, Q Yang, S FuProceedings of the 4th ACM/IEEE Symposium on Edge Computing, 2019.dl.acm.org, pp. 88–100. Available at: <https://doi.org/10.1145/3318216.3363300>.

- Chou, F.-C., et al., 2020. Predicting motion of vulnerable road users using high-definition maps and efficient convnets. In: 2020 IEEE Intelligent Vehicles Symposium (IV) IEEE 1655–1662.
- Choudhury, A., et al., 2016. An integrated V2X simulator with applications in vehicle platooning. In: In 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). IEEE, pp. 1017–1022.
- Chuma, E.L., Iano, Y., 2020. A movement detection system using continuous-wave Doppler radar sensor and convolutional neural network to detect cough and other gestures. *IEEE Sens. J.* 21 (3), 2921–2928.
- Cui, H. and Dahoun, N. (2021) ‘High precision human detection and tracking using millimeter-wave radars’, *IEEE Aerospace and Electronic Systems Magazine*, 36(1), pp. 22–32. Available at: <https://doi.org/10.1109/MAES.2020.3021322>.
- Cui, Y., et al. (2022) ‘Deep Learning for Image and Point Cloud Fusion in Autonomous Driving: A Review’, *IEEE Transactions on Intelligent Transportation Systems*. Institute of Electrical and Electronics Engineers Inc., pp. 722–739. Available at: <https://doi.org/10.1109/TITS.2020.3023541>.
- Dabral, S., et al., 2014. Trends in camera based automotive driver assistance systems (adas). In: 2014 IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS) IEEE 1110–1115.
- Dandil, E., Çevik, K.K., 2019. Computer vision based distance measurement system using stereo camera view. In: In 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMISIT). IEEE, pp. 1–4.
- Dao, T.-T., Pham, Q.-V., Hwang, W.-J., 2022. FastMDE: a fast CNN architecture for monocular depth estimation at high resolution. *IEEE Access* 10, 16111–16122.
- Darugna, F., et al. (2019) ‘RTK and PPP-RTK using smartphones: From short-baseline to long-baseline applications’, in Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019), pp. 3932–3945.
- Dayangac, E., et al., 2016. Target position and speed estimation using lidar. In: In International Conference on Image Analysis and Recognition. Springer, pp. 470–477.
- Demkowicz, J., 2022. Non-least square GNSS positioning algorithm for densely urbanized areas. *Remote Sens. (Basel)* 14 (9), 2027.
- Dinh-Van, N., et al., 2017. Indoor Intelligent Vehicle localization using WiFi received signal strength indicator. in: 2017 IEEE MTT-S international conference on microwaves for intelligent mobility (ICMIM) IEEE 33–36.
- Dong, X., et al., 2021. Radar Camera Fusion via Representation Learning in Autonomous Driving. In: In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1672–1681.
- Du, Y., et al., 2021. Vulnerabilities and integrity of precise point positioning for intelligent transport systems: overview and analysis. *Satellite Navigation* 2 (1), 1–22.
- Du, Y., et al. (2022) ‘Quantifying the performance and optimizing the placement of roadside sensors for cooperative vehicle-infrastructure systems’, *IET Intelligent Transport Systems*, 16(7), pp. 908–925. Available at: <https://doi.org/10.1049/ITR2.12185>.
- Eckelmann, S., et al., 2017. V2v-communication, lidar system and positioning sensors for future fusion algorithms in connected vehicles. *Transp. Res. Procedia* 27, 69–76.
- El Bouziady, A., et al., 2018. Vehicle speed estimation using extracted SURF features from stereo images. In: In 2018 International Conference on Intelligent Systems and Computer Vision (ISCV). IEEE, pp. 1–6.
- Emara, M., Filippou, M.C., Sabella, D., 2020. MEC-enhanced information freshness for safety-critical C-V2X communications. In: In 2020 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, pp. 1–5.
- Feng, D., et al., 2019. Deep active learning for efficient training of a lidar 3d object detector. in: 2019 IEEE Intelligent Vehicles Symposium (IV) IEEE 667–674.
- Flores, C., et al. (2019) ‘A Cooperative Car-Following/Emergency Braking System with Prediction-Based Pedestrian Avoidance Capabilities’, *IEEE Transactions on Intelligent Transportation Systems*, 20(5), pp. 1837–1846. Available at: <https://doi.org/10.1109/TITS.2018.2841644>.
- Frank, R., Hawlader, F., 2021. Poster: commercial 5G performance: a V2X experiment. in: 2021 IEEE Vehicular Networking Conference (VNC) IEEE 129–130.
- Fujikami, S., et al. (2015) ‘Fast Device Discovery for Vehicle-to-Pedestrian communication using wireless LAN’, in: 2015 12th Annual IEEE Consumer Communications and Networking Conference, CCNC 2015. Institute of Electrical and Electronics Engineers Inc., pp. 35–40. Available at: <https://doi.org/10.1109/CCNC.2015.7157943>.
- Gao, X., et al., 2021. Perception through 2D-MIMO FMCW automotive radar under adverse weather. In: In 2021 IEEE International Conference on Autonomous Systems (ICAS). IEEE, pp. 1–5.
- Gao, H., et al. (2018) ‘Object Classification Using CNN-Based Fusion of Vision and LIDAR in Autonomous Vehicle Environment’, *IEEE Transactions on Industrial Informatics*, 14(9), pp. 4224–4230. Available at: <https://doi.org/10.1109/TII.2018.2822828>.
- Garcia, E., et al., 2015. A robust UWB indoor positioning system for highly complex environments. In: In 2015 IEEE International Conference on Industrial Technology (ICIT). IEEE, pp. 3386–3391.
- Garcia, M.H.C. et al. (2021) ‘A Tutorial on 5G NR V2X Communications’, *IEEE Communications Surveys & Tutorials*, 23(3), pp. 1972–2026. Available at: <https://doi.org/10.1109/COMST.2021.3057017>.
- Gelbal, S.Y., Aksun-Guvenc, B., and Guvenc, L. (2020) ‘Collision Avoidance of Low Speed Autonomous Shuttles with Pedestrians’, *International Journal of Automotive Technology*, 21(4), pp. 903–917. Available at: <https://doi.org/10.1007/s12239-020-0087-7>.
- Gelbal, S.Y. et al. (2017) ‘Elastic band based pedestrian collision avoidance using V2X communication’, in: 2017 IEEE Intelligent Vehicles Symposium (IV), pp. 270–276. Available at: <https://doi.org/10.1109/IVS.2017.77995731>.
- Ghosh, A. et al. (2019) ‘5G Evolution: A View on 5G Cellular Technology Beyond 3GPP Release 15’, *IEEE Access*, 7, pp. 127639–127651. Available at: <https://doi.org/10.1109/ACCESS.2019.2939938>.
- Gigl, T., et al., 2007. Analysis of a UWB indoor positioning system based on received signal strength. in: 2007 4th Workshop on Positioning Navigation and Communication. IEEE 97–101.
- Global road safety statistics | Brake (2018). Available at: <https://www.brake.org.uk/get-involved/take-action/mybrake/knowledge-centre/global-road-safety> (Accessed: 19 August 2022).
- GMV (2011) ‘RTK Standards - Navipedia’. Available at: https://gssc.esa.int/navipedia/index.php/RTK_Standards.
- Gómez-Huélamo, C., et al. (2021) ‘SmartMOT: Exploiting the fusion of HDMaps and Multi-Object Tracking for Real-Time scene understanding in Intelligent Vehicles applications’, in: 2021 IEEE Intelligent Vehicles Symposium (IV), pp. 710–715. Available at: <https://doi.org/10.1109/IV48863.2021.9575443>.
- González-Martín, M., et al. (2019) ‘Analytical Models of the Performance of C-V2X Mode 4 Vehicular Communications’, *IEEE Transactions on Vehicular Technology*, 68(2), pp. 1155–1166. Available at: <https://doi.org/10.1109/TVT.2018.2888704>.
- Guo, X., et al., 2018. Pedestrian detection based on fusion of millimeter wave radar and vision. In: In Proceedings of the 2018 International Conference on Artificial Intelligence and Pattern Recognition, pp. 38–42.
- Haimovich, A.M., Blum, R.S. and Cimini, L.J. (2008) ‘MIMO Radar with Widely Separated Antennas’, *IEEE Signal Processing Magazine*, 25(1), pp. 116–129. Available at: <https://doi.org/10.1109/MSP.2008.4408448>.
- Han, T.Y., Song, B.C., 2016. Night vision pedestrian detection based on adaptive preprocessing using near infrared camera. In: In 2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia). IEEE, pp. 1–3.
- He, S., Li, J. and Qiu, T.Z. (2017) ‘Vehicle-to-Pedestrian Communication Modeling and Collision Avoiding Method in Connected Vehicle Environment’, *Transportation Research Record: Journal of the Transportation Research Board*, 2621(1), pp. 21–30. Available at: <https://doi.org/10.3141/2621-03>.
- Hellmers, H., et al., 2013. An IMU/magnetometer-based indoor positioning system using Kalman filtering. In: In International Conference on Indoor Positioning and Indoor Navigation. IEEE, pp. 1–9.
- Heyman, J., 2019. TracTrac: A fast multi-object tracking algorithm for motion estimation. *Comput. Geosci.* 128, 11–18.
- Hoang, G.-M., et al., 2017. Robust data fusion for cooperative vehicular localization in tunnels. in: 2017 IEEE Intelligent Vehicles Symposium (IV) IEEE 1372–1377.
- Hou, X., Wang, Y., Chau, L.-P., 2019. Vehicle tracking using deep sort with low confidence track filtering. In: In 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, pp. 1–6.
- Hu, H., et al. (2021) ‘Investigating the impact of multi-lidar placement on object detection for autonomous driving’, openaccess.thecvf.com [Preprint]. Available at: http://openaccess.thecvf.com/content/CVPR2022/html/Hu_Investigating_the_Impact_of_Multi-LiDAR_Placement_on_Object_Detection_for_CVPR_2022_paper.html (Accessed: 27 September 2023).
- Huang, J., et al. (2020) ‘Recent advances and challenges in security and privacy for V2X communications’, *IEEE Open Journal of Vehicular Technology*, 1, pp. 244–266. Available at: <https://doi.org/10.1109/OJVT.2020.2999885>.
- Hulse, L.M., Xie, H. and Galea, E.R. (2018) ‘Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age’, *Safety Science*, 102, pp. 1–13. Available at: <https://doi.org/10.1016/j.ssci.2017.10.001>.
- Hung, C.-M., et al., 2019. 9.1 toward automotive surround-view radars. In: In 2019 IEEE International Solid-State Circuits Conference-(ISSCC). IEEE, pp. 162–164.
- Hussein, A., et al., 2016. P2V and V2P communication for pedestrian warning on the basis of autonomous vehicles. In: In 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). IEEE, pp. 2034–2039.
- Hyun, E., Jin, Y.-S., Lee, J.-H., 2016. A pedestrian detection scheme using a coherent phase difference method based on 2D range-Doppler FMCW radar. *Sensors* 16 (1), 124.
- Immoreev, I.I. and Fedotov, P.G.S.D. V (2002) ‘Ultra wideband radar systems: advantages and disadvantages’, in: 2002 IEEE Conference on Ultra Wideband Systems and Technologies (IEEE Cat. No. 02EX580). IEEE, pp. 201–205.
- Jakowski, N., Stankov, S.M., Klaehn, D., 2005. Operational space weather service for GNSS precise positioning. *Ann. Geophys. Copernicus GmbH* 3071–3079.
- Jeong, S., Baek, Y., Son, S.H., 2016. A hybrid V2X system for safety-critical applications in VANET. in: 2016 IEEE 4th international conference on cyber-physical systems, networks, and applications (CPSNA) IEEE 13–18.
- Jianyong, Z., et al., 2014. RSSI based Bluetooth low energy indoor positioning. In: In 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE, pp. 526–533.
- Jung, C., et al., 2020. V2X-communication-aided autonomous driving: system design and experimental validation. *Sensors* 20 (10), 2903.
- Karaiim, M., et al., 2018. GNSS error sources. *Multifunctional Operation and Application of GPS* 69–85.
- Karoui, M., Freitas, A. and Chalhoub, G. (2020) ‘Performance comparison between LTE-V2X and ITS-G5 under realistic urban scenarios’, in: 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), pp. 1–7. Available at: <https://doi.org/10.1109/VTC2020-Spring48590.2020.9129423>.
- Kawanishi, T., et al. (2019) ‘Simple Secondary Radar for Non-Line-of-Sight Pedestrian Detection’, in: 2019 IEEE Conference on Antenna Measurements & Applications (CAMA), pp. 151–152. Available at: <https://doi.org/10.1109/CAMA47423.2019.8959735>.
- Kiel, K., et al. (2020) ‘Review of V2X-IoT Standards and Frameworks for ITS Applications’, *Applied Sciences*, 10(12), p. 4314. Available at: <https://doi.org/10.3390/app10124314>.

- Kim, T. and Park, T. (2019) 'Placement optimization of multiple lidar sensors for autonomous vehicles', ieeexplore.ieee.org/abstract/document/8718317/ [Preprint]. Available at: <https://ieeexplore.ieee.org/abstract/document/8718317/> (Accessed: 28 September 2023).
- Kim, J.H., Batchuluun, G., Park, K.R., 2018. Pedestrian detection based on faster R-CNN in nighttime by fusing deep convolutional features of successive images. *Expert Syst. Appl.* 114, 15–33.
- Kim, J., Kim, J., Cho, J., 2019. An advanced object classification strategy using YOLO through camera and LiDAR sensor fusion. In: In 2019, 13th International Conference on Signal Processing and Communication Systems, ICSPCS 2019 - Proceedings. Institute of Electrical and Electronics Engineers Inc., Available at. <https://doi.org/10.1109/ICSPCS47537.2019.9008742>.
- Kim, G. et al. (2020) 'MulRan: Multimodal Range Dataset for Urban Place Recognition', in Proceedings - IEEE International Conference on Robotics and Automation. Institute of Electrical and Electronics Engineers Inc., pp. 6246–6253. Available at: <https://doi.org/10.1109/ICRA40945.2020.9197298>.
- Kong, H., et al., 2019. OBU design and test analysis with centimeter-level positioning for LTE-V2X. In: In 2019 5th International Conference on Transportation Information and Safety (ICTIS). IEEE, pp. 383–387.
- Kuo, C.-H., Lin, C.-C., Sun, J.-S., 2017. Modified microstrip Franklin array antenna for automotive short-range radar application in blind spot information system. *IEEE Antennas Wirel. Propag. Lett.* 16, 1731–1734.
- Kuo, Y.-S. et al. (2014) 'Luxapose: Indoor positioning with mobile phones and visible light', in Proceedings of the 20th annual international conference on Mobile computing and networking, pp. 447–458.
- Kwon, S.K., et al., 2017. Detection scheme for a partially occluded pedestrian based on occluded depth in lidar-radar sensor fusion. *Opt. Eng.* 56 (11), 113112.
- Kwon, S.K. et al. (2016) 'A low-complexity scheme for partially occluded pedestrian detection using LiDAR-radar sensor fusion', in 2016 IEEE 22nd International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA). IEEE, p. 104.
- Lahmyed, R., El Ansari, M., Ellahyani, A., 2019. A new thermal infrared and visible spectrum images-based pedestrian detection system. *Multimed. Tools Appl.* 78 (12), 15861–15885.
- Lang, Y., et al., 2020. Person identification with limited training data using radar micro-Doppler signatures. *Microw. Opt. Technol. Lett.* 62 (3), 1060–1068.
- Lee, J., et al., 2022. GAN-based LiDAR translation between sunny and adverse weather for autonomous driving and driving simulation. *Sensors* 22 (14), 5287. <https://doi.org/10.3390/s22145287>.
- Lee, G.H., Kwon, K.H. and Kim, M.Y. (2019) 'Ambient Environment Recognition Algorithm Fusing Vision and LiDAR Sensors for Robust Multi-channel V2X System', in 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN), pp. 98–101. Available at: <https://doi.org/10.1109/ICUFN.2019.8806087>.
- Lee, S., Kim, D., 2016. An energy efficient vehicle to pedestrian communication method for safety applications. *Wirel. Pers. Commun.* 86 (4), 1845–1856. <https://doi.org/10.1007/s11277-015-3160-1>.
- Lee, J.-E. et al. (2017) 'Harmonic clutter recognition and suppression for automotive radar sensors', *International Journal of Distributed Sensor Networks*, 13(9), p. 1550147717729793.
- Lee, G.H. et al. (2020) 'Object Detection Using Vision and LiDAR Sensor Fusion for Multi-channel V2X System', in 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), pp. 1–5. Available at: <https://doi.org/10.1109/ICAIIIC48513.2020.9065243>.
- Lekic, V., Babic, Z., 2019. Automatic radar and camera fusion using generative adversarial networks. *Comput. Vis. Image Underst.* 184, 1–8.
- Lekidis, A., Bouali, F., 2021. C-V2X network slicing framework for 5G-enabled vehicle platooning applications. in 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring) IEEE 1–7.
- Li, G., et al., 2019. Novel 4D 79 GHz Radar Concept for Object Detection and Active Safety Applications. In: GeMIC 2019–2019 German Microwave Conference. Institute of Electrical and Electronics Engineers Inc., pp. 87–90.
- Li, P., Chen, X. and Shen, S. (2019) 'Stereo r-cnn based 3d object detection for autonomous driving', in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7644–7652.
- Li, Y., Ibanez-Guzman, J., 2020. Lidar for autonomous driving: the principles, challenges, and trends for automotive lidar and perception systems. *IEEE Signal Process Mag.* 37 (4), 50–61. <https://doi.org/10.1109/MSP.2020.2973615>.
- Li, J., Stoica, P., 2008. MIMO radar signal processing. John Wiley & Sons.
- Li, X., Wang, J., Liu, C., 2015. A Bluetooth/PDR integration algorithm for an indoor positioning system. *Sensors* 15 (10), 24862–24885.
- Li, C.Y. et al. (2018) 'V2PSense: Enabling Cellular-Based V2P Collision Warning Service through Mobile Sensing', in IEEE International Conference on Communications. Institute of Electrical and Electronics Engineers Inc. Available at: <https://doi.org/10.1109/ICC.2018.8422981>.
- Li, W. et al. (2020) 'Smot: Single-shot multi object tracking', arXiv preprint arXiv: 2010.16031 [Preprint].
- Lianghai, J., et al., 2018. Multi-RATs support to improve V2X communication. in 2018 IEEE wireless communications and networking conference (WCNC) IEEE 1–6.
- Liu, G., et al., 2017. A blind spot detection and warning system based on millimeter wave radar for driver assistance. *Optik* 135, 353–365.
- Liu, H., 2020. 'Autonomous rail rapid transit (ART) systems', in robot systems for rail transit applications. Elsevier 189–234. <https://doi.org/10.1016/b978-0-12-822968-2.00005-x>.
- Liu, Q., et al., 2021. A V2X-integrated positioning methodology in ultradense networks. *IEEE Internet Things J.* 8 (23), 17014–17028. <https://doi.org/10.1109/JIOT.2021.3075532>.
- Liu, Z., et al., 2021. Robust target recognition and tracking of self-driving cars with radar and camera information fusion under severe weather conditions. *IEEE Transactions on Intelligent Transportation Systems*.
- Liu, W., Muramatsu, S. and Okubo, Y. (2018) 'Cooperation of V2I/P2I communication and roadside radar perception for the safety of vulnerable road users', in Proceedings of 2018 16th International Conference on Intelligent Transport System Telecommunications, ITST 2018. Institute of Electrical and Electronics Engineers Inc. Available at: <https://doi.org/10.1109/ITST.2018.8566704>.
- Liu, B., Kamel, A.E., 2016. V2X-based decentralized cooperative adaptive cruise control in the vicinity of intersections. *IEEE Trans. Intell. Transp. Syst.* 17 (3), 644–658. <https://doi.org/10.1109/TITS.2015.2486140>.
- Liu, Q.i., Liang, P., et al., 2021. A highly accurate positioning solution for C-V2X systems. *Sensors* 21 (4), 1175.
- Liu, Q., Liang, P., Xia, J., Wang, T., Song, M., Xu, X., Zhang, J., Fan, Y., Liu, L., 2021. A highly accurate positioning solution for C-V2X systems. *Sensors* 21 (4), 1175.
- Liu, Q.i., Song, M., et al., 2021. High Accuracy Positioning for C-V2X. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing, p. 012100.
- Liu, K., Wang, W., Wang, J., 2019. Pedestrian detection with lidar point clouds based on single template matching. *Electronics* 8 (7), 780. <https://doi.org/10.3390/electronics8070780>.
- Liu, Zishan et al. (2016) 'Implementation and performance measurement of a V2X communication system for vehicle and pedestrian safety', *International Journal of Distributed Sensor Networks*, 12(9), p. 1550147716671267.
- Long, Y. et al. (2021) 'Radar-camera pixel depth association for depth completion', in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12507–12516.
- Ma, S., et al., 2019. An efficient V2X based vehicle localization using single RSU and single receiver. *IEEE Access* 7, 46114–46121.
- Machardy, Z., et al., 2018. V2X access technologies: Regulation, research, and remaining challenges. *IEEE Commun. Surv. Tutorials* 20 (3), 1858–1877. <https://doi.org/10.1109/COMST.2018.2808444>.
- Mafakheri, B., et al., 2021. Optimizations for hardware-in-the-loop-based v2x validation platforms. in 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring) IEEE 1–7.
- Malik, R.Q., et al., 2020. An overview on V2P communication system: Architecture and application. In: In 2020 3rd International Conference on Engineering Technology and Its Applications, ICETEA 2020. Institute of Electrical and Electronics Engineers Inc., pp. 174–178. <https://doi.org/10.1109/ICETEA50496.2020.9318863>
- Mansouri, A., Martinez, V. and Härrí, J. (2019) 'A First Investigation of Congestion Control for LTE-V2X Mode 4', in 2019 15th Annual Conference on Wireless On-demand Network Systems and Services (WONS), pp. 56–63. Available at: <https://doi.org/10.23919/WONS.2019.8795500>.
- Maruta, K. et al. (2021) 'Blind-Spot Visualization via AR Glasses using Millimeter-Wave V2X for Safe Driving', in 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall), pp. 1–5. Available at: <https://doi.org/10.1109/VTC2021-Fall52928.2021.9625498>.
- Meyer, M., Kuschk, G., 2019. 'Automotive radar dataset for deep learning based 3d object detection', in 2019 16th european radar conference (EuRAD). IEEE 129–132.
- Meyer, M., Kuschk, G., 2019. 'Deep learning based 3d object detection for automotive radar and camera', in 2019 16th European Radar Conference (EuRAD). IEEE 133–136.
- Milz, S. et al. (2018) 'Visual slam for automated driving: Exploring the applications of deep learning', in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 247–257.
- Ming, Y., et al., 2021. Deep learning for monocular depth estimation: A review. *Neurocomputing* 438, 14–33. <https://doi.org/10.1016/j.neucom.2020.12.089>.
- Miucic, R. et al. (2018) 'V2X Applications Using Collaborative Perception', in 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), pp. 1–6. Available at: <https://doi.org/10.1109/VTCFall.2018.8690818>.
- Molina-Masegosa, R., Gozalvez, J., 2017. *LTE-V for sidelink 5G V2X vehicular communications: A new 5G technology for short-range vehicle-to-everything communications*. *IEEE Veh. Technol. Mag.* 12 (4), 30–39.
- Muhammad, M., Safdar, G.A., 2018. Survey on existing authentication issues for cellular-assisted V2X communication. *Veh. Commun.* 12, 50–65. <https://doi.org/10.1016/j.vehcom.2018.01.008>.
- Mumuni, F., Mumuni, A., 2022. Bayesian cue integration of structure from motion and CNN-based monocular depth estimation for autonomous robot navigation. *International Journal of Intelligent Robotics and Applications* 6 (2), 191–206. <https://doi.org/10.1007/s41315-022-00226-2>.
- Murphy, Y.L., et al., 2018. Accurate pedestrian path prediction using neural networks. In: In 2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings. Institute of Electrical and Electronics Engineers Inc., pp. 1–7. <https://doi.org/10.1109/SSCI.2017.8285398>
- Musha, H. and Fujii, M. (2017) 'A study on indoor positioning based on RTK-GPS', in 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE). IEEE, pp. 1–2.
- Nabati, R. and Qi, H. (2021) 'Centerfusion: Center-based radar and camera fusion for 3d object detection', in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 1527–1536.
- Naiden, A. et al. (2019) 'Shift r-cnn: Deep monocular 3d object detection with closed-form geometric constraints', in 2019 IEEE International Conference on Image Processing (ICIP). IEEE, pp. 61–65.
- Najman, P., Zemčík, P., 2020. Vehicle speed measurement using stereo camera pair. *IEEE Transactions on Intelligent Transportation Systems* [preprint].

- Napolitano, A., et al. (2019). 'Implementation of a MEC-based vulnerable road user warning system', in 2019 AEIT international conference of electrical and electronic technologies for automotive (AEIT AUTOMOTIVE). IEEE 1–6.
- Naranjo, J.E. et al. (2017) 'Application of vehicle to another entity (V2X) communications for motorcycle crash avoidance', *Journal of Intelligent Transportation Systems*, 21(4), pp. 285–295. Available at: <https://doi.org/10.1080/15472450.2016.1247703>.
- Nardini, G., et al. (2018). Cellular-V2X communications for platooning: Design and evaluation. *Sensors* 18 (5), 1527.
- Ni, Y. et al. (2020) 'A V2X-based Approach for Avoiding Potential Blind-zone Collisions between Right-turning Vehicles and Pedestrians at Intersections', in 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), pp. 1–6. Available at: <https://doi.org/10.1109/ITSC45102.2020.9294501>.
- Nielsen, T.A.S. and Haustein, S. (2018) 'On sceptics and enthusiasts: What are the expectations towards self-driving cars?', *Transport Policy*, 66, pp. 49–55. Available at: <https://doi.org/10.1016/j.tranpol.2018.03.004>.
- Ninnemann, J. et al. (2022) 'Multipath-Assisted Radio Sensing and State Detection for the Connected Aircraft Cabin', *Sensors*, 22(8). Available at: <https://doi.org/10.3390/s22082859>.
- Okuda, R., Kajiwara, Y., Terashima, K., 2014. 'A survey of technical trend of ADAS and autonomous driving', in Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test, VLSI-DAT 2014. IEEE Computer Society. <https://doi.org/10.1109/VLSI-DAT.2014.6834940>.
- Palfy, A., et al. (2020). CNN based road user detection using the 3D radar cube. *IEEE Rob. Autom. Lett.* 5 (2), 1263–1270.
- Palfy, A., Kooij, J.F.P. and Gavrila, D.M. (2019) 'Occlusion aware sensor fusion for early crossing pedestrian detection', in 2019 IEEE Intelligent Vehicles Symposium (IV), pp. 1768–1774. Available at: <https://doi.org/10.1109/IVS.2019.8814065>.
- Palfy, A. et al. (2022) 'Multi-Class Road User Detection with 3+1D Radar in the View-of-Delft Dataset', *IEEE Robotics and Automation Letters*, 7(2), pp. 4961–4968. Available at: <https://doi.org/10.1109/LRA.2022.3147324>.
- Parada, R., et al., 2021. Machine learning-based trajectory prediction for VRU collision avoidance in V2X environments. in 2021 IEEE Global Communications Conference (GLOBECOM) IEEE 1–6.
- Pearre, N.S., Ribberink, H., 2019. Review of research on V2X technologies, strategies, and operations. *Renew. Sustain. Energy Rev.* 105, 61–70. <https://doi.org/10.1016/j.rser.2019.01.047>.
- Pérez, R., et al., 2018. 'Single-frame vulnerable road users classification with a 77 GHz FMCW radar sensor and a convolutional neural network', in 2018 19th International Radar Symposium (IRS). IEEE 1–10.
- Postica, G., Romanoni, A., Matteucci, M., 2016. Robust moving objects detection in lidar data exploiting visual cues. In: In 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 1093–1098.
- Qi, B., et al., 2016. Pedestrian detection from thermal images: A sparse representation based approach. *Infrared Phys. Technol.* 76, 157–167.
- Qin, T., et al., 2021. A Light-Weight Semantic Map for Visual Localization towards Autonomous Driving. In: In 2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 11248–11254.
- Quenzel, J. and Behrke, S. (2021) 'Real-time Multi-Adaptive-Resolution-Surfel 6D LiDAR Odometry using Continuous-time Trajectory Optimization', in IEEE International Conference on Intelligent Robots and Systems. Institute of Electrical and Electronics Engineers Inc., pp. 5499–5506. Available at: <https://doi.org/10.1109/IROS51168.2021.9636763>.
- Radjou, A.N. and Kumar, S.M. (2018) 'Epidemiological and clinical profile of fatality in vulnerable road users at a high volume trauma center', *Journal of Emergencies, Trauma and Shock*, 11(4), pp. 282–287. Available at: https://doi.org/10.4103/JETS.JETS_55_17.
- Rahman, M.L., et al., 2019. Framework for a perceptive mobile network using joint communication and radar sensing. *IEEE Trans. Aerosp. Electron. Syst.* 56 (3), 1926–1941.
- Rauch, A. et al. (2012) 'Car2x-based perception in a high-level fusion architecture for cooperative perception systems', in 2012 IEEE Intelligent Vehicles Symposium. Available at: <https://ieeexplore.ieee.org/abstract/document/6232130/> (Accessed: 28 September 2023).
- Rawashdeh, Z. and Wang, Z. (2018) 'Collaborative automated driving: A machine learning-based method to enhance the accuracy of shared information', 21st International Conference on Intelligent Transportation Systems (ITSC) [Preprint]. Available at: <https://ieeexplore.ieee.org/abstract/document/8569832/> (Accessed: 28 September 2023).
- Rebut, J., et al., 2022. Raw High-Definition Radar for Multi-Task Learning. In: In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 17021–17030.
- Ren, J., et al., 2019. Information fusion of digital camera and radar. in 2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC) IEEE 1–4.
- Repala, V.K., Dubey, S.R., 2019. Dual CNN models for unsupervised monocular depth estimation. In: In International Conference on Pattern Recognition and Machine Intelligence. Springer, pp. 209–217.
- Rizos, C. et al. (2012) 'Precise Point Positioning: Is the era of differential GNSS positioning drawing to an end?'. Road casualties Great Britain: e-Scooter (2021). Available at: <https://www.gov.uk/government/statistics/reported-road-casualties-great-britain-e-scooter-factsheet-2021/reported-road-casualties-great-britain-e-scooter-factsheet-2021-provisional> (Accessed: 19 August 2022).
- Road casualties in Great Britain: pedal cycle (2020). Available at: <https://www.gov.uk/government/statistics/reported-road-casualties-great-britain-pedal-cyclist-factsheet-2020/reported-road-casualties-great-britain-pedal-cyclist-factsheet-2020#what-type-of-road> (Accessed: 19 August 2022).
- Road casualty statistics in Great Britain (2017). Available at: <https://maps.dft.gov.uk/road-casualties/index.html> (Accessed: 19 August 2022).
- Roriz, R. et al. (2022) 'DIOR: A Hardware-Assisted Weather Denoising Solution for LiDAR Point Clouds', *IEEE Sensors Journal*, 22(2), pp. 1621–1628. Available at: <https://doi.org/10.1109/JSEN.2021.3133873>.
- Rujikigmatjorn, S., Watcharapinchai, N., 2017. Real-time hog-based pedestrian detection in thermal images for an embedded system. In: In 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, pp. 1–6.
- Ruß, T., Krause, J. and Schönrock, R. (2016) 'V2X-based cooperative protection system for vulnerable road users and its impact on traffic', in ITS World Congress.
- Saleem, M.K., et al., 2017. Lens antenna for wide angle beam scanning at 79 GHz for automotive short range radar applications. *IEEE Trans. Antennas Propag.* 65 (4), 2041–2046.
- Salman, Y.D., Ku-Mahamud, K.R., Kamioka, E., 2017. Distance measurement for self-driving cars using stereo camera. *International Conference on Computing and Informatics 235–242*.
- Sander, U., Lubbe, N., Pietzsch, S., 2019. Intersection AEB implementation strategies for left turn across path crashes. *Traffic Inj. Prev.* 20 (sup1), S119–S125.
- Schumann, O., et al., 2021. RadarScenes: A real-world radar point cloud data set for automotive applications. In: In 2021 IEEE 24th International Conference on Information Fusion (FUSION). IEEE, pp. 1–8.
- Segata, M., Arvani, P., Cigno, R.L., 2021. A critical assessment of C-V2X resource allocation scheme for platooning applications. In: In 2021 16th Annual Conference on Wireless on-Demand Network Systems and Services Conference (WONS). IEEE, pp. 1–8.
- Sempere-García, D., Sepulcre, M., Gozalvez, J., 2021. LTE-V2X Mode 3 scheduling based on adaptive spatial reuse of radio resources. *Ad Hoc Netw.* 113, 102351.
- Sengupta, A., et al., 2020. mm-Pose: Real-time human skeletal posture estimation using mmWave radars and CNNs. *IEEE Sens. J.* 20 (17), 10032–10044.
- Sheeny, M. et al. (2021) 'Radiate: A Radar Dataset for Automotive Perception in Bad Weather', in Proceedings - IEEE International Conference on Robotics and Automation. Institute of Electrical and Electronics Engineers Inc., pp. 5617–5623. Available at: <https://doi.org/10.1109/ICRA48506.2021.9562089>.
- Shen, C., et al., 2020. Seamless GPS/inertial navigation system based on self-learning square-root cubature Kalman filter. *IEEE Trans. Ind. Electron.* 68 (1), 499–508.
- Shi, F. et al. (2022) 'Pi-NIC: Indoor Sensing Using Synchronized Off-The-Shelf Wireless Network Interface Cards and Raspberry Pis', in 2022 2nd IEEE International Symposium on Joint Communications & Sensing (JC&S), pp. 1–6. Available at: <https://doi.org/10.1109/JCSS54387.2022.9743512>.
- Shrestha, R. et al. (2020) 'Evolution of V2X Communication and Integration of Blockchain for Security Enhancements', *Electronics*, 9(9), p. 1338. Available at: <https://doi.org/10.3390/electronics9091338>.
- Smith, G.M. (2021) Types of ADAS Sensors in Use Today | Dewesoft. Available at: <https://dewesoft.com/daq/types-of-adas-sensors#types> (Accessed: 1 September 2022).
- Song, W., et al., 2020. CNN-based 3D object classification using Hough space of LiDAR point clouds. *HCIS* 10 (1), 1–14.
- Stephenson, S. et al. (2012) 'Implementation of V2X with the integration of Network RTK: Challenges and solutions', in Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012), pp. 1556–1567.
- Talbot, R. et al. (2017) 'Fatal and serious collisions involving pedal cyclists and trucks in London between 2007 and 2011', *Traffic Injury Prevention*, 18(6), pp. 657–665. Available at: <https://doi.org/10.1080/15389588.2017.1291938>.
- Tang, Z., et al., 2018. Single-camera and inter-camera vehicle tracking and 3D speed estimation based on fusion of visual and semantic features. In: In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 108–115.
- Thomä, R., et al., 2021. Joint communication and radar sensing: An overview. In: In 2021 15th European Conference on Antennas and Propagation (EuCAP). IEEE, pp. 1–5.
- Thompson, A.W., 2018. Economic implications of lithium ion battery degradation for Vehicle-to-Grid (V2G) services. *J. Power Sources* 396, 691–709.
- Ting, S.L., et al., 2011. The study on using passive RFID tags for indoor positioning. *International Journal of Engineering Business Management* 3, 8.
- Toker, O., Alsweiss, S., 2020. MmWave Radar Based Approach for Pedestrian Identification in Autonomous Vehicles. In: In Conference Proceedings - IEEE SOUTHEASTCON. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/SoutheastCon44009.2020.9249704>.
- Toledo-Moreo, R., et al., 2018. Positioning and digital maps. *Intelligent Vehicles*. Elsevier 141–174.
- Tong, X., et al., 2019. A double-step unscented Kalman filter and HMM-based zero-velocity update for pedestrian dead reckoning using MEMS sensors. *IEEE Trans. Ind. Electron.* 67 (1), 581–591.
- Toulminet, G., et al., 2006. Vehicle detection by means of stereo vision-based obstacles features extraction and monocular pattern analysis. *IEEE Trans. Image Process.* 15 (8), 2364–2375.
- Tripathi, N. and Yogamani, S. (2020) 'Trained Trajectory based Automated Parking System using Visual SLAM on Surround View Cameras', arXiv preprint arXiv: 2001.02161 [Preprint].
- Protecting Vulnerable Road Users (VRU) With V2P Tech - AUTOCRYPT (2022). Available at: <https://autocrypt.io/protecting-vru-with-v2p-technology/#:~:text=Vulnerable>

- %20road%20user%20(VRU)%20is,or%20someone%20in%20a%20wheelchair.
(Accessed: 31 August 2022).
- Vargas Rivero, J.R. et al. (2020) 'Weather Classification Using an Automotive LIDAR Sensor Based on Detections on Asphalt and Atmosphere', *Sensors*, 20(15), p. 4306. Available at: <https://doi.org/10.3390/s20154306>.
- Vázquez-Gallego, F. et al. (2019) 'Demo: A Mobile Edge Computing-based Collision Avoidance System for Future Vehicular Networks', in IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 904–905. Available at: <https://doi.org/10.1109/INFOMW.2019.8845107>.
- Wagner, J. et al. (2016) 'Multispectral Pedestrian Detection using Deep Fusion Convolutional Neural Networks.', in *ESANN*, pp. 509–514.
- Walters, J.G., et al., 2019. Rural Positioning Challenges for Connected and Autonomous Vehicles. In: In Proceedings of the 2019 International Technical Meeting of the Institute of Navigation, pp. 828–842.
- Wang, J., et al., 2016. A tightly-coupled GPS/INS/UWB cooperative positioning sensors system supported by V2I communication. *Sensors* 16 (7), 944.
- Wang, J., et al., 2019. A survey of vehicle to everything (V2X) testing. *Sensors* 19 (2), 334.
- Wang, L., et al., 2020c. 'High dimensional frustum pointnet for 3D object detection from camera, LiDAR, and radar', in 2020 IEEE Intelligent Vehicles Symposium (IV). IEEE 1621–1628.
- Wang, W., Sakurada, K. and Kawaguchi, N. (2017) 'Reflectance Intensity Assisted Automatic and Accurate Extrinsic Calibration of 3D LiDAR and Panoramic Camera Using a Printed Chessboard', *Remote Sensing*, 9(8), p. 851. Available at: <https://doi.org/10.3390/rs9080851>.
- Wang, B. et al. (2020) Fusion Positioning System Based on IMU and Roadside LiDAR in Tunnel for C-V2X Use. SAE Technical Paper.
- Wang, T.H. et al. (2020) 'V2VNet: Vehicle-to-Vehicle Communication for Joint Perception and Prediction', Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12347 LNCS, pp. 605–621. Available at: https://doi.org/10.1007/978-3-030-58536-5_36.
- Warren ME (2019) 'Automotive LiDAR technology', in 2019 Symposium on VLSI Circuits. Available at: <https://ieeexplore.ieee.org/abstract/document/8777993/> (Accessed: 4 October 2023).
- Weber, R., Misener, J., Park, V., 2019. 'C-V2X - A Communication Technology for Cooperative, Connected and Automated Mobility', in Mobile Communication - Technologies and Applications; 24. ITG-Symposium 1–6.
- Wen, L.-H., Jo, K.-H., 2021. Fast and accurate 3D object detection for lidar-camera-based autonomous vehicles using one shared voxel-based backbone. *IEEE Access* 9, 22080–22089.
- Wicaksono, D.W., Setiyono, B., 2017. Speed estimation on moving vehicle based on digital image processing. *IJCSAM (International Journal of Computing Science and Applied Mathematics)* 3 (1), 21–26.
- Widmann, G.R. et al. (2000) 'Comparison of lidar-based and radar-based adaptive cruise control systems', JSTOR [Preprint]. Available at: <https://www.jstor.org/stable/44699119> (Accessed: 4 October 2023).
- Wild, T., Braun, V. and Viswanathan, H. (2021) 'Joint Design of Communication and Sensing for Beyond 5G and 6G Systems', *IEEE Access*, 9, pp. 30845–30857. Available at: <https://doi.org/10.1109/ACCESS.2021.3059488>.
- Wu, R., et al., 2019. 'Modified driving safety field based on trajectory prediction model for pedestrian-vehicle collision', *Sustainability* (Switzerland), 11(22). Available at: <https://doi.org/10.3390/su11226254>.
- Wu, Q., et al., 2020c. Hybrid SVM-CNN classification technique for human–vehicle targets in an automotive LFM CW radar. *Sensors* 20 (12), 3504.
- Wu, X. et al. (2013) 'Vehicular Communications Using DSRC: Challenges, Enhancements, and Evolution', *IEEE Journal on Selected Areas in Communications*, 31(9), pp. 399–408. Available at: <https://doi.org/10.1109/JSAC.2013.SUP.0513036>.
- Wu, J. et al. (2020) 'Vehicle Detection under Adverse Weather from Roadside LiDAR Data', *Sensors*, 20(12), p. 3433. Available at: <https://doi.org/10.3390/s20123433>.
- Wu, Q. et al. (2020) 'Performance Analysis of Cooperative Intersection Collision Avoidance with C-V2X Communications', in 2020 IEEE 20th International Conference on Communication Technology (ICCT), pp. 757–762. Available at: <https://doi.org/10.1109/ICCT50939.2020.9295949>.
- Wu, T. et al. (2021) 'A Pedestrian Detection Algorithm Based on Score Fusion for Multi-LiDAR Systems', *Sensors*, 21(4), p. 1159. Available at: <https://doi.org/10.3390/s21041159>.
- Wymeersch, H. et al. (2021) 'Integration of Communication and Sensing in 6G: a Joint Industrial and Academic Perspective', in 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 1–7. Available at: <https://doi.org/10.1109/PIMRC50174.2021.9569364>.
- Xiao, Z., et al., 2019. A unified multiple-target positioning framework for intelligent connected vehicles. *Sensors* 19 (9), 1967.
- Xu, W., et al., 2015. Indoor positioning for multiphotodiode device using visible-light communications. *IEEE Photonics J.* 8 (1), 1–11.
- Xu, R. et al. (2022) 'Opv2v: An open benchmark dataset and fusion pipeline for perception with vehicle-to-vehicle communication', in International Conference on Robotics and Automation (ICRA). Available at: <https://ieeexplore.ieee.org/abstract/document/9812038/> (Accessed: 27 September 2023).
- Xu, R. et al. (2022) 'V2X-ViT: Vehicle-to-Everything Cooperative Perception with Vision Transformer', Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 13699 LNCS, pp. 107–124. Available at: https://doi.org/10.1007/978-3-031-19842-7_7.
- Xu, Runsheng et al. (2023) 'V2v4real: A real-world large-scale dataset for vehicle-to-vehicle cooperative perception', in Conference on Computer Vision and Pattern Recognition (CVPR). Available at: http://openaccess.thecvf.com/content_CVPR2023/html/Xu_V2V4Real_A_Real-World_Large-Scale_Dataset_for_Vehicle-to-Vehicle_Cooperative_Perception_CVPR_2023_paper.html (Accessed: 28 September 2023).
- Yadav, A.K., Szpytko, J., 2017. Safety problems in vehicles with adaptive cruise control system. *Journal of KONBIN* 42 (1), 389.
- Yamazato, T. (2015) 'Image sensor based visible light communication for V2X', in 2015 IEEE Summer Topicals Meeting Series (SUM), pp. 165–166. Available at: <https://doi.org/10.1109/PHOSST.2015.7248248>.
- Yao, L., et al., 2017. An integrated IMU and UWB sensor based indoor positioning system. In: In 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE, pp. 1–8.
- Yasir, M., Ho, S.-W., Vellambi, B.N., 2014. Indoor positioning system using visible light and accelerometer. *J. Lightwave Technol.* 32 (19), 3306–3316.
- Yazici, A., Yayan, U., Yuclu, H., 2011. 'An ultrasonic based indoor positioning system', in 2011 International Symposium on Innovations in Intelligent Systems and Applications. IEEE 585–589.
- Ye, M. et al. (2020) HVNet: Hybrid Voxel Network for LiDAR Based 3D Object Detection.
- Yi, C., Zhang, K. and Peng, N. (2019) 'A multi-sensor fusion and object tracking algorithm for self-driving vehicles', Proceedings of the Institution of Mechanical Engineers, Part D: Journal of automobile engineering, 233(9), pp. 2293–2300.
- Yin, X. et al. (2014) 'Performance and Reliability Evaluation of BSM Broadcasting in DSRC with Multi-Channel Schemes', *IEEE Transactions on Computers*, 63(12), pp. 3101–3113. Available at: <https://doi.org/10.1109/TC.2013.175>.
- Yoshioka, M. et al. (2018) 'Real-time object classification for autonomous vehicle using LiDAR', in ICIIIBMS 2017 - 2nd International Conference on Intelligent Informatics and Biomedical Sciences. Institute of Electrical and Electronics Engineers Inc., pp. 210–211. Available at: <https://doi.org/10.1109/ICIIIBMS.2017.8279696>.
- Yu, H., et al., 2022. Dair-v2x: A large-scale dataset for vehicle-infrastructure cooperative 3d object detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [preprint]. Available at.
- Yucel, H., et al., 2012. 'Development of indoor positioning system with ultrasonic and infrared signals', in 2012 International Symposium on Innovations in Intelligent Systems and Applications. IEEE 1–4.
- Yusuf, S.A., Aldawsari, A.A. and Souissi, R. (2022) 'Automotive parts assessment: applying real-time instance-segmentation models to identify vehicle parts', arXiv preprint arXiv:2202.00884 [Preprint].
- Zaarane, A., et al., 2020. Distance measurement system for autonomous vehicles using stereo camera. *Array* 5, 100016.
- Zakuan, F.R.A., et al., 2017. Threat assessment algorithm for active blind spot assist system using short range radar sensor. *ARPN Journal of Engineering and Applied Sciences* 12, 4270–4275.
- Zhang, J., et al., 2020. Vehicle tracking and speed estimation from roadside lidar. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 13, 5597–5608.
- Zhang, Z. (2000) 'A flexible new technique for camera calibration', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), pp. 1330–1334. Available at: <https://doi.org/10.1109/34.888718>.
- Zhang, J.A. et al. (2017) 'Framework for an Innovative Perceptive Mobile Network Using Joint Communication and Sensing', in 2017 IEEE 85th Vehicular Technology Conference (VTC Spring), pp. 1–5. Available at: <https://doi.org/10.1109/VTCSPRING.2017.8108564>.
- Zhang, A. et al. (2021) 'Perceptive Mobile Networks: Cellular Networks With Radio Vision via Joint Communication and Radar Sensing', *IEEE Vehicular Technology Magazine*, 16(2), pp. 20–30. Available at: <https://doi.org/10.1109/MVT.2020.3037430>.
- Zhao, Y., et al., 2008. Implementing indoor positioning system via ZigBee devices. In: In 2008 42nd Asilomar Conference on Signals, Systems and Computers. IEEE, pp. 1867–1871.
- Zhou, Y. and Tuzel, O. (2018) 'VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection', in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, pp. 4490–4499. Available at: <https://doi.org/10.1109/CVPR.2018.00472>.
- Zhou, Y. et al. (2019) 'End-to-End Multi-View Fusion for 3D Object Detection in LiDAR Point Clouds'. Available at: <http://arxiv.org/abs/1910.06528> (Accessed: 19 August 2022).
- Zhou, H. et al. (2020) 'Evolutionary V2X Technologies Toward the Internet of Vehicles: Challenges and Opportunities', Proceedings of the IEEE, 108(2), pp. 308–323. Available at: <https://doi.org/10.1109/JPROC.2019.2961937>.
- Zhou, Y. et al. (2022) 'Towards Deep Radar Perception for Autonomous Driving: Datasets, Methods, and Challenges', *Sensors*, 22(11), p. 4208. Available at: <https://doi.org/10.3390/s22114208>.
- Zhuang, Y. et al. (2022) 'Illumination and Temperature-Aware Multispectral Networks for Edge-Computing-Enabled Pedestrian Detection', *IEEE Transactions on Network Science and Engineering*, 9(3), pp. 1282–1295. Available at: <https://doi.org/10.1109/TNSE.2021.3139335>.
- Zoghloomi, C., Kacimi, R., Dhaou, R., 2022. 'A Study on Dynamic Collection of Cooperative Awareness Messages in V2X Safety Applications', in 2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC). IEEE 723–724.