Autonomous Vehicle, Sensing and Communication Survey

Igal Bilik,
School of Electrical and Computer Engineering
Ben Gurion University of the Negev, Israel

Abstract

Autonomous driving is revolutionizing public transportation and affects our daily lives. However, we currently witness some decrease in the number of newly established companies introducing autonomous vehicles. Beyond the post-COVID global economic situation, this report analyzes the technological showstoppers and proposes possible research directions to imitate these obstacles. It focuses on studying the research directions that could enable a robust and reliable autonomous vehicle operation in practical urban conditions. Considering the practical challenges of the autonomous vehicle cost, operation in complex and diverse scenarios, adverse weather conditions, and regulation challenges, this report summarises the last year literature survey. It proposes a few possible directions for the near-future research.

I. Introuction

Road traffic accidents caused approximately $1,35$ million deaths worldwide in 2016. Active driver assistance systems (ADAS) have been shown to dramatically reduce traffic accidents and causalities. The current global ADAS market of $\$27$ billion is expected to grow to $\$83$ billion by 2030. Multiple active safety features and some level of autonomy have already been introduced in the majority of commercial vehicles, and the automotive industry is progressing toward autonomous driving. An autonomous vehicle substitutes human drivers in both sensing and decision-making. Therefore, it needs to collect information on the host's vehicle location and its surroundings. The situation awareness sensors are used to provide reliable, dense, and timely information on the vehicle's surroundings. In particular, they need to acquire information on the vehicle's drivable path and all objects above the road level. Sensors need to detect, localize, and classify objects that could interfere with autonomous driving.

Autonomous vehicles are typically equipped with multiple sensors of multiple modalities, such as cameras, radars, and LiDARs [1]. Cameras resemble human driver vision and are the most natural sensors for autonomous driving. However, cameras are sensitive to adverse weather and poor lighting conditions, do not provide native range and velocity measurements, and have to be mounted behind the optically transparent fascia. Therefore, the automotive sensing suit typically includes radars [2]–[4]. In turn, radars are robust to adverse weather conditions, insensitive to lighting conditions, provide long and accurate range measurements, and can be packaged behind optically nontransparent fascia [5].

In recent years, autonomous vehicles have become a reality in multiple publically deployed autonomous vehicle fleets. However, their wide deployment requires improved reliability for the positive acceptance [6]. The short-, medium-, and long-term effects of autonomous vehicles on urban transportation and the environment were conducted in [7]. It was concluded that autonomous vehicles expect to influence urban transportation and human mobility by reducing vehicle ownership, public and active travel, traffic delay and congestion, and travel costs and increasing accessibility, mobility, vehicle miles traveled, and revenue generation for commercial operators. The long-term effects include encouraging dispersed urban development, reducing parking demand, and enhancing network capacity. Autonomous vehicles will reduce energy consumption and protect the environment by reducing greenhouse gas emissions. They also expect to reduce traffic crashes involving human errors and increase the convenience and productivity of passengers by facilitating multitasking. However, most people are concerned about personal safety, security, and privacy. The literature also shows that shared autonomous vehicles are well-positioned to significantly positively impact transportation and the urban environment more than private autonomous vehicles. shared autonomous vehicles in a dynamic ride-sharing situation could be an effective policy option to reduce vehicle ownership, traffic congestion, and travel time and improve the overall performance of the transportation system. Researchers proposed to formulate appropriate funding mechanisms and policies to encourage ride-sharing and on-demand mobility among travelers to increase the use of shared autonomous vehicles. Thus, pertinent policies in transportation (e.g., automation of transit, integration of transit and non-motorized transport, encouraging shared and micro-mobility), infrastructure (e.g., adjustment and redesign of existing roads), and urban planning (e.g., update of urban development plans, land-use plans, parking policies, and design, green belts) are essential to realizing the benefits of autonomous vehicles. Moreover, the law and order situation needs to be improved to provide safety and security to passengers while sharing AVs.

This report summarises the major gaps in the autonomous vehicles technology reported in the literature during 2023.

II. Misbehavior Detection

Connected and autonomous vehicles have various embedded components connected through different communication technologies, and their security has become a vital concern [8]. Therefore, Misbehavior Detection plays a significant role in enabling vehicles to quickly identify security risks and adopt effective immediate countermeasures.

Misbehavior detection for autonomous vehicle communications refers to the ability of self-driving cars to identify and respond to potentially problematic or malicious behavior in their communication with other vehicles and systems [9]. This can include detecting and responding to attempts to jam or interfere with their communication signals [10], identifying and responding to false or misleading information, and recognizing and responding to attempts to hack or compromise their systems. Misbehavior detection is important for maintaining the security and reliability of the vehicle's communication systems, and for ensuring the safe and efficient operation of self-driving cars. It typically involves the use of advanced algorithms and machine learning techniques to monitor and analyze the vehicle's communication with other systems and to identify and respond to potential misbehavior.

There is a need to develop approaches for misbehavior detection and the resulting cybersecurity on various levels:

- Detection of the intended or unintended sensing interferer for radars, cameras, and LiDARs.
- Secured communication between vehicles (V2V) and vehicle to cloud (V2C).
- Secured central computers and faults-tolerant processing.

III. Adverse Weather

However, perception and sensing for autonomous driving under adverse weather conditions have been the problem that kept autonomous vehicles (AVs) from going to higher autonomy for a long time [11].

Weather phenomena have various negative influences. On average, global precipitation occurs 11.0% of the time [12]. It has been proven that the risk of accidents in rain conditions is 70% higher than normal. 77% of the countries in the world receive snow. The United States national statistics show that each year over $30,000$ vehicle crashes occur on snowy or icy roads or during snowfall or sleet. Phenomena like fog, haze, sandstorms, and strong light severely decrease visibility and raise driving risks [13]. Secondary problems directly or circumstantially caused by weather, such as heat coldness, and contamination, also have unpredictable or undesirable effects on autonomous vehicles.

Harsh weather affects all sensors used for autonomous vehicles, radars, cameras, and LiDARs [14]. However, cameras and LiDARs are more sensitive to weather conditions. therefore, the development of the weather-immune sensing suit is critical for the practical deployment of autonomous vehicles.

Addressing harsh weather conditions requires the development of weather-immune sensing capabilities for LiDARs, cameras, and radars. The research necessary involves sensing technology, signal processing algorithms, and fusion between multiple sensing modalities.

IV. Vehicle Control

Vehicle control is one of the most critical challenges in autonomous vehicles and connected and automated vehicles are paramount in vehicle safety, passenger comfort, transportation efficiency, and energy saving [15]. Nevertheless, certain critical parameters like longitudinal and lateral velocity, sideslip angle, orientation, and tire forces cannot be directly gauged in the context of commercial vehicles. These pivotal metrics can only be inferred indirectly. Within this realm, the sideslip angle holds particular significance as it encompasses both longitudinal and lateral velocity details, making its accuracy a key indicator of the precision of estimations for longitudinal and lateral velocity. Despite its importance, the sideslip angle proves to be the most intricate among the mentioned parameters, primarily due to the influence of errors from other states, such as longitudinal velocity and roll angle [16]. Additionally, the sideslip angle plays a crucial role in vehicle stabilization, motion planning, road condition assessment, handover modules, and vehicle navigation [17]. In essence, without an accurately determined sideslip angle, the performance of the aforementioned applications may be compromised or even lead to inevitable failure. Consequently, the efficient and robust estimation of the vehicle sideslip angle has garnered significant attention in the vehicle control community over the past few years.

While other states, such as longitudinal speed information, are also imperative for vehicle control applications, their estimation is comparably simpler, and ample literature has effectively addressed these concerns. Thus, in this endeavor, our focus is directed towards the estimation of the sideslip angle. Given its pivotal role in vehicle control, addressing this sideslip angle estimation problem becomes imperative, necessitating a comprehensive review for readers. Drawing from existing literature, we explore three primary approaches for sideslip angle estimation, leveraging accessible data from affordable onboard sensors, including the steering wheel angle sensor (SWAS), wheel speed sensors (WSS), the inertial measurement unit (IMU), and global navigation satellite systems (GNSS) [18]. These approaches include the onboard-sensor-based (OSB) approach, GNSS-augmented (GAU) approach, and artificial-intelligence-augmented (AIA) approach [19].

The following are the possible proposed research topics to address the autonomous vehicle control challenge: **Multi-modal state estimation.**
Beyond relying solely on GNSS, incorporating LiDAR and cameras into the mix can enhance vehicle state estimation, thereby boosting overall accuracy. Nevertheless, these additional sensors are vulnerable to environmental influences, including factors such as buildings, lighting conditions, and weather. Moreover, their sampling latency and frequency exhibit substantial variations. Consequently, it becomes imperative to assess the real-time confidence level of sensor signals and devise robust fusion algorithms to mitigate these challenges.

Robust trajectory tracking control with state uncertainty
The presence of multi-modal information enables the real-time estimation of three-dimensional vehicle attitude information. This expansion allows the vehicle control model to transition from a two-dimensional plane to a three-dimensional space. Given that state estimation results may exhibit a degree of uncertainty, it becomes paramount to consider the influence of this uncertainty on the algorithm's performance when crafting trajectory control algorithms.

Control challenge of heterogeneous dynamics in mixed-autonomy vehicle platoons
Addressing the collaboration and coordination between human-driven and autonomous vehicles poses a substantial challenge for collaborative control techniques in mixed-autonomy traffic scenarios. A key obstacle lies in the diverse dynamics introduced by the behavior of human drivers within the vehicle platoon. Surmounting this challenge necessitates the development of innovative control techniques capable of adapting to the variability of human drivers, all while upholding safety and efficiency on the road. Potential solutions may involve the integration of machine learning algorithms for predicting human driver behavior and the implementation of hybrid control architectures that amalgamate centralized and decentralized control techniques. This approach enables AVs to adjust to the conduct of human-driven vehicles while maintaining

coordination within the platoon.

Resilience control techniques in the presence of cyber-attacks and communication failures

In a dynamic and unpredictable driving environment, the occurrence of communication and sensing failures and potential disruptions by adversarial entities is a realistic concern. Consequently, it becomes imperative to incorporate resilience control techniques into the collaborative control framework. This integration empowers the system to persistently adapt to evolving conditions, even when faced with faults, errors, or malicious attacks. Drawing upon insights from the domain of cyber-physical systems proves to be an effective strategy for crafting resilient collaborative control techniques for Connected and Autonomous Vehicles. Moreover, it is crucial to factor in security and privacy requirements. Designing communication, sensing, and control protocols with security and privacy considerations, which may involve encryption, authentication, and access control mechanisms, becomes essential for ensuring the future safety and performance of Connected and Autonomous Vehicles.

V. Computing System Architecture

The computing system is crucial in Intelligent Vehicles, ensuring safety, security, energy efficiency, and effective communication. Given that a typical autonomous vehicle is equipped with an array of onboard sensors, such as Lidar, cameras, Radar, communication modules, and the global navigation satellite system (GNSS), the generated data per minute can be substantial. To manage the efficient processing and fusion of heterogeneous information, the computing system must be adept at real-time information processing to maintain safe Autonomous Driving.

In [20], it was shown that the two primary computing architectures are widely employed in Intelligent Vehicles: the modular design method and the end-to-end design method. The modular design approach involves decoupling functional units into separate modules facilitating system implementation, fault diagnosis, and module updates. Following this approach, the computing architecture of Intelligent Vehicles can be categorized into critical modules, including computation, communication, storage, security and privacy, and power management. Conversely, contemporary artificial intelligence strongly influences end-to-end computing architecture, relying predominantly on learning-based approaches to directly process sensing data and generate control outputs.

Despite notable advancements in the computing systems of Intelligent Vehicles in recent years, several critical constraints persist across hardware, software, sensor layers, and more, hindering the widespread deployment of commercial autonomous vehicles. A comprehensive analysis in [21] identifies five key design constraints, namely performance constraints, predictability constraints, storage constraints, thermal constraints, and power constraints for Intelligent Vehicles.

Performance constraints primarily manifest at the application level. While Intelligent Vehicles perform fundamental functions such as perception, planning, decision-making, and control, a performance gap remains compared to human capabilities. Key performance constraints include frame rate and processing latency. Given that a human driver reacts within $100 - 150$ ms, the vehicle, for safety reasons, should respond even faster, ideally within a latency of 100 ms and a frequency of at least once every 100 ms [21].

Predictability is another critical constraint, encompassing both temporal aspects (meeting time deadlines) and functional aspects (making correct decisions). In addition to conventional mean latency metrics, tail latency (99.99th percentile latency) is advocated to meet stringent predictability requirements [21].

Storage constraints emerge as an essential bottleneck for energy conservation and computing performance on Intelligent Vehicles. A single autonomous vehicle can generate $2-40$ TB of data daily, necessitating high-speed, voluminous storage space. Real-time data storage and transfer also significantly boost energy and power usage.

Thermal constraints pose considerations in two aspects: maintaining the operational temperature range and limiting the impact of heat generated by the computing system on the vehicle's thermal profile. Without a cooling system, the computing system's 1 kW power consumption can increase the in-cabin temperature by up to 10^o per minute. Implementing the computing system in a climate-controlled area with an additional cooling system becomes necessary to mitigate thermal impacts.

Power constraints, a critical aspect influencing autonomous vehicle capabilities, encompass power consumption of the computing system, storage consumption, and cooling overhead. The cumulative effect of computing consumption and additional energy consumption for storage and cooling can significantly decrease the mileage of Intelligent Vehicles, particularly electric ones. For instance, a high-energy system like GPU can reduce fuel efficiency by up to 11.5% [21]. Therefore, there is a pressing need for more energy-efficient computing systems and green AI techniques in the future of Intelligent Vehicles.

VI. Explainable Artificial Intelligence

Autonomous vehicles already have a significant commercial, ethical, and technological effect [22]. The development of intelligent vehicles is the most promising path for driving autonomy. Modern autonomous vehicles use artificial intelligence for a variety of tasks. While pipeline architectures break down the driving task into more manageable components, providing a somewhat interpretable processing of sensor data through specialized modules (perception, planning, decision, control), they have several drawbacks. Firstly, they rely on human heuristics and manually selected intermediate representations, which lack proof of optimality for the driving task. Secondly, their handcrafted nature limits their ability to accommodate real-world uncertainties, hindering their generalization to scenarios not anticipated by system designers. Furthermore, from an engineering perspective, these systems are challenging to scale and maintain due to the interdependency of various modules [23]. Lastly, they are susceptible to error propagation among the multiple sub-modules [24].

Autonomous driving is a safety-critical application. Therefore, the performance guarantees are required. However, artificial intelligence (AI) based self-driving vehicles are not entirely testable under all scenarios as it is not possible to exhaustively list and evaluate every situation they may encounter. The AI-based solutions overcome some of the limitations of the modular pipeline stack. They are sometimes described as black boxes for their critical lack of transparency and interpretability. Indeed, as trained within the deep learning paradigm, they fall into the known shortcomings associated with these architectures. In this paper, we focus on models learned by behavior cloning, which leverages datasets of human driving sessions, as opposed to reinforcement learning approaches, which train models through trial-and-error simulation. As a fallback solution, this motivates the need for an explanation of driving decisions. The onboard explainability of perception, localization, decision-making, and control, grounded in multimodal signals from sensor systems, is a crucial design prerequisite for the future deployment and global acceptance of commercial autonomous vehicles [25]. The explainability of autonomous vehicles can be further dissected into interpretability and completeness. Within interpretability, transparency and post-hoc explanation emerge as two primary branches [26]. Additionally, achieving local explanation versus global explanation necessitates distinct computing system designs for autonomous vehicles [26]. Essentially, the design criteria for an explainable interface for autonomous vehicles align with societal and legal requirements and play a pivotal role in fostering safer, transparent, publicly accepted, and environmentally friendly autonomous vehicles [27].

VII. Artificial Intelligence for Automotive Radar Processing

Modern sensing suit includes LiDARs, cameras, and radars, where radars provide extended object detection ranges, direct measurement of an object's velocity, and reliability in adverse weather and poor lighting conditions [1], [3], [28], [29].

Radars obtain information on the vehicle surroundings by processing electromagnetic echoes returned from the reflective surfaces [30]. Radars' operation in dense urban environments is characterized by reflections from multiple artificial surfaces, such as buildings and guardrails [31], [32]. Automotive radars operate at the 77GHz frequency band, with a short wavelength of ~ 3.8mm [1], [30]. Therefore, these reflective surfaces are flat compared with the radar wavelength and, as a result, induce multipath propagation phenomena [33]–[36].

The indirect radar echoes returned via reflection from the auxiliary flat surfaces are frequently indistinguishable from the direct target reflections [33]. When direct and indirect reflections occupy different range-Doppler-direction cells, the conventional radar receiver processes them separately. As a result, the indirect reflections may be interpreted as "ghost" targets [33]. These "ghost" targets degrade radar performance by increasing the probability of false alarms, decreasing the probability of detection, and resulting in a significant waste of radar computational resources [30].

Accurate modeling of the multipath propagation phenomena in practical urban scenarios is challenged by the large number of reflective surfaces in the scene and the variety of their properties. Moreover, short radar wavelength and the lack of millimeter-level host vehicle localization accuracy in dynamic automotive scenarios induce time-varying clutter that is received at the non-zero Doppler frequencies. These conditions limit the applicability of the conventional ray-tracing-based multipath propagation modeling approaches, which are widely used in cellular networks [37]–[40].

Recently, data-driven deep learning (DL) approaches addressing the high model dimensionality, non-linearity, and modeling complexity challenges were introduced for automotive radars' target detection [41]–[45], classification [46]–[52], tracking [50], [53], and interference mitigation [54]–[61]. The DL approaches require an extensive, accurately annotated database for deep neural network (DNN) training. Data annotation is an extremely expensive and labor-intensive task [62]–

[64]. Multiple semi-automatic and learning-based approaches were proposed for the annotation of automotive visual [65]–[68] and LiDAR [69], [70] datasets. Recently, several approaches for annotation burden reduction by combining the synthetic with the collected data were proposed in the literature [71], [72].

Radar data annotation is even more challenging for human annotators because its appearance is non-intuitive and unnatural compared to visual data [70], [73]–[75]. Furthermore, annotating the automotive radar data collected in the multipath-dominated urban scenarios is even more labor-demanding, requires expert-level knowledge, and is prone to multiple annotation errors associated with the "ghost" targets appearance [76].

Recently, the automotive radar data annotation approach using auxiliary sensors was introduced in [77]. This approach exploited the correlation between radar and LiDAR detections to annotate the automotive radar data. However, LiDAR's inherent limitations, such as limited range and sensing sparsity, can result in false- and miss-labeling [78]. Moreover, this approach does not allow fine-granularity labeling to distinguish between the actual target and the reflector, the reflector type, and the multipath reflection order.

Therefore, there is a need for radar multipath annotation to address the problem in [79] as one of the major challenges in DNN-based automotive radar processing. One possible research direction is the development of automated annotation methods to convert the publically available datasets into fine-granularity multipath-annotated datasets, with the labeling level that is currently available only in the small limited datasets [76].

VIII. Detection using Diffusion Models for Autonomous Driving

The development of diffusion models [80] has shown to be remarkably successful in image generation, surpassing other approaches' performance. Diffusion is also a better approach for generating traffic scenes. Latent diffusion is especially well suited for generating traffic scenes due to decoupling latent and output spaces [81], [82].

A probabilistic model based on diffusion, referred to here as a "diffusion model" for conciseness, is a parameterized Markov chain trained through variational inference to generate samples that align with the data over a finite period. The transitions within this chain are learned to invert a diffusion process, which involves a Markov chain progressively introducing noise to the data in the reverse direction of sampling until the signal is obliterated. In cases where the diffusion involves minor increments of Gaussian noise, it is feasible to configure the transitions of the sampling chain as conditional Gaussians, enabling a notably straightforward neural network parameterization [83]. There is a need for further research to exploit all potential of diffusion models for autonomous driving.

IX. Vision Language Models for Autonomous Driving

Closed-set 3D perception models trained on only a pre-defined set of object categories can be inadequate for safety-critical applications such as autonomous driving, where new object types can be encountered after deployment [84]. Ove the recent few years, a novel learning paradigm, Pre-training, Fine-tuning, and Prediction, has showcased significant effectiveness across a broad spectrum of visual recognition tasks [85]. Within this paradigm, a Deep Neural Network (DNN) model undergoes an initial phase of pre-training using off-the-shelf large-scale training data, whether annotated or unannotated [86]. Following this, the pre-trained model is fine-tuned using task-specific annotated training data. Harnessing the comprehensive knowledge acquired during pre-training, this learning paradigm expedites network convergence and facilitates the training of high-performing models for various downstream tasks.

However, despite its success, the Pre-training, Fine-tuning, and Prediction paradigm still requires an additional stage of task-specific fine-tuning using labeled training data from each downstream task. Drawing inspiration from advancements in natural language processing [87], a novel deep learning paradigm known as Vision-Language Model Pre-training and Zero-shot Prediction has recently garnered increased attention [88]. In this paradigm, a Vision-Language Model (VLM) undergoes pre-training using large-scale image-text pairs abundantly available online [89]. The pre-trained VLM can then be directly applied to downstream visual recognition tasks without fine-tuning. The VLM pre-training is typically guided by specific vision-language objectives that enable the learning of image-text correspondences from extensive image-text pairs [90], [91]. Exploiting these paradigms is a promising area to be investigated in future research for a variety of autonomous driving tasks.

X. Few-short Learning for Autonomous Driving

Autonomous vehicles encounter a multitude of unknown objects, some of which are rarely seen or entirely new, such as the next generation of concept cars or traffic signs adorned with graffiti [92]. Addressing this challenge necessitates the development of innovative approaches to object recognition problems that involve less human supervision and fewer annotated datasets.

While various methods attempt to utilize web images for training deep learning models, the issue of noisy images persists, and the reliance on human-supplied search keywords remains a form of human supervision. Two solutions, namely One-Shot Learning (OSL) and Few-Shot Learning (FSL), offer the capability to learn new categories with just one or a few images, respectively [93], [94], [95], [92]. Few-shot learning (FSL) involves learning new classes using a minimal training dataset, typically consisting of one or a few images per category. FSL is intricately connected to knowledge transfer, wherein a model previously trained on extensive data is applied to a similar task with reduced training data. The effectiveness of FSL's generalization hinges on the accuracy of the transferred knowledge. Furthermore, numerous approaches utilize meta-learning to comprehend the challenges of few-shot or few-example learning. The primary hurdle lies in enhancing generalization ability, as FSL frequently grapples with the issue of overfitting. In the case of one-shot learning, there is only 1 example per class in the supporting set, thus it faces more challenge in comparison to the FSL and Zero-shot learning is the extreme case of the FSL.

XI. Generative AI for Autonomous Driving

The prospective solution of the vehicular mixed reality (MR) Metaverse for achieving autonomous driving involves integrating physical and virtual vehicular networks [96], [97]. Enabling multi-dimensional communications among physical and virtual entities can eliminate the isolation of "data islands" on roads, leading to enhanced road safety and traffic efficiency while concurrently lowering energy consumption and carbon emissions [98]. With the support of digital twin (DT) technologies, autonomous vehicles (AVs) leverage advanced sensors such as ultrasonic radars, cameras, and LiDAR to gather data from their surroundings and construct virtual representations in the virtual space [99]. Subsequently, AVs can make driving decisions, including driving model selection and motion planning, utilizing artificial intelligence (AI) methods. Despite the presence of panoramic cameras and high-grade LiDAR in AVs, each vehicle can only capture limited environmental data and may miss aspects like occlusions []. Therefore, the collaboration of multiple connected AVs, roadside units (RSUs), and virtual simulators becomes crucial for sharing and merging sensing data in the virtual space [100]. This collaboration enables the perception of comprehensive environmental information, including occlusions [101]. However, directly collecting realistic driving data on a large scale to train AVs in the physical world is challenging and expensive. Addressing these challenges requires research in generative AI [101].

XII. Semantic Segmentation for Autonomous Driving

While deep learning has found extensive application in computer vision, particularly in areas like self-driving cars, the emphasis has often been on achieving state-of-the-art results without sufficient attention to the safety implications of different classes. In many models, all classes are treated uniformly, and the focus is primarily on average precision. However, not all classes contribute equally to the reliability and safety of autonomous driving systems. For instance, the Person class should be accorded higher priority than the Sky class in terms of segmentation accuracy [102].

A critical aspect for future research involves addressing how semantic segmentation for autonomous driving systems can be enhanced by considering salient class considerations. This encompasses dealing with challenging instances of crucial classes and maintaining a balanced trade-off between accuracy and inference latency.

The imperative is to elevate semantic segmentation for autonomous driving systems by giving precedence to the precise detection and segmentation of critical classes, such as persons, which directly impact safety. Additionally, there is a need to tackle instances where important classes may be partially obscured or manifest in different appearances.

XIII. Communication

The decision-making capacity of individual vehicles is constrained by onboard computing and storage resources, posing challenges in coping with the demands of dense and mixed traffic flow environments. Consequently, autonomous vehicles must enhance their intelligence levels and broaden sensing and decision-making capabilities by leveraging external information and resources through wireless communications.

Unlike traditional cellular network mobile terminals, vehicles' high-speed movement leads to dynamic changes in network topology and frequent communication link switches [103]. Simultaneously, the intricate and variable driving environment introduces multipath effects and interference from other signals for vehicular communication. Safety-critical applications, like collision avoidance and platooning of Intelligent Vehicles (autonomous vehicle), necessitate networks to meet stringent low latency and ultrahigh reliability requirements, often challenging for traditional wireless networks. Therefore, autonomous vehicle should adopt dedicated communication technologies to ensure efficient and stable interactions between vehicles, infrastructure, and cloud platforms [104].

As an extension of the perception capabilities of Autonomous Driving (AD) and autonomous vehicles, vehicular communication involves the vehicles themselves and various elements of transportation, communication, and other systems [105]. This integration and convergence represent a focal point for the automotive, transportation, communication, and information industries [106]. Vehicular communication technology originated from academic research and demonstration projects in Europe and the United States at the end of the 20th century, initially known as vehicular ad-hoc network (VANET) [107]. With advancements in information and communication technologies, the concept evolved from VANET to the Internet of Vehicles, gaining attention for its supporting role in intelligent transportation and AD [108].

Expanding on this concept, vehicular communication technology has progressed from VANET connecting vehicles and infrastructure to Vehicle-to-Everything (V2X) [109], linking various transportation system elements, such as vehicles, infrastructures, pedestrians, and clouds [110]. Specifically, V2X encompasses two types of technologies: 1) dedicated short-range communication (DSRC) standardized by IEEE and 2) cellular V2X (C-V2X) standardized by the 3rd Generation Partnership Project (3GPP). In terms of communication methods, V2X communication can be categorized into V2V, V2I, vehicle-to-pedestrian (V2P), and vehicle-to-cloud/network (V2C/N) communication. Consequently, vehicular networks facilitate real-time and efficient information interaction among pedestrians, vehicles, infrastructures, and clouds, supporting AD and autonomous vehicles' massive data transmission needs.

The further development of connected and shared autonomous vehicles requires standardization of the communication protocols, security of the vehicle-to-cloud communication, and adoption of the higher-speed communication approaches.

XIV. Proposed Research Topics

- Detection of the intended or unintended sensing interferer for radars, cameras, and LiDARs.
- Secured communication between vehicles (V2V) and vehicle to cloud (V2C).
- Secured central computers and faults-tolerant processing.
- Harsh weather detection and sensing in adverse weather conditions.
- Robust and secured autonomous vehicle control .
- Low consumption power and on-the-edge compute architectures.
- Explainable artificial intelligence for autonomous driving.
- Artificial intelligence for automotive radar processing.

- Diffusion models for autonomous driving.
- Vision language models for autonomous driving.
- Few-short learning for autonomous driving.
- Generative artificial intelligence for autonomous driving.
- Advanced semantic segmentation for autonomous driving.
- Vehicle-to-vehicle and Vehicle-to-Cloud resilient communication and standardization.

REFERENCES

[1] I. Bilik, "Comparative analysis of radar and lidar technologies for automotive applications," *IEEE Intelligent Transportation Systems Magazine*, vol. 15, no. 1, pp. 244–269, 2023.

[2] F. Engels, P. Heidenreich, M. Wintermantel, L. Stäcker, M. Al Kadi, and A. M. Zoubir, "Automotive radar signal processing: Research directions and practical challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 865–878, 2021.

[3] I. Bilik, O. Longman, S. Villeval, and J. Tabrikian, "The rise of radar for autonomous vehicles: Signal processing solutions and future research directions," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 20–31, 2019.

[4] S. Saponara, M. S. Greco, and F. Gini, "Radar-on-chip/in-package in autonomous driving vehicles and intelligent transport systems: Opportunities and challenges," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 71–84, 2019.

[5] F. Gini, "Grand challenges in radar signal processing," in *Frontiers in Signal Processing*, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:234281562

[6] K. Othman, "Public attitude towards autonomous vehicles before and after crashes: A detailed analysis based on the demographic characteristics," *Cogent Engineering*, vol. 10, no. 1, p. 2156063, 2023. [Online]. Available: https://doi.org/10.1080/23311916.2022.2156063

[7] M. M. Rahman and J.-C. Thill, "Impacts of connected and autonomous vehicles on urban transportation and environment: A comprehensive review," *Sustainable Cities and Society*, vol. 96, p. 104649, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670723002603

[8] M. L. Bouchouia, H. Labiod, O. Jelassi, J.-P. Monteuuis, W. B. Jaballah, J. Petit, and Z. Zhang, "A survey on misbehavior detection for connected and autonomous vehicles," *Vehicular Communications*, vol. 41, p. 100586, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214209623000165

[9] R. W. van der Heijden, T. Lukaseder, and F. Kargl, "Veremi: A dataset for comparable evaluation of misbehavior detection in vanets," in *Security and Privacy in Communication Networks*, R. Beyah, B. Chang, Y. Li, and S. Zhu, Eds. Cham: Springer International Publishing, 2018, pp. 318–337.

[10] K. M. Alheeti, A. Gruebler, and K. McDonald-Maier, "Using discriminant analysis to detect intrusions in external communication for self-driving vehicles," *Digital Communications and Networks*, vol. 3, no. 3, pp. 180–187, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352864817300913

[11] Y. Zhang, A. Carballo, H. Yang, and K. Takeda, "Perception and sensing for autonomous vehicles under adverse weather conditions: A survey," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 196, pp. 146–177, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271622003367

[12] K. E. Trenberth and Y. Zhang, "How often does it really rain?" *Bulletin of the American Meteorological Society*, vol. 99, no. 2, pp. 289 – 298, 2018. [Online]. Available: https://journals.ametsoc.org/view/journals/bams/99/2/bams-d-17-0107.1.xml

[13] A. Mehra, M. Mandal, P. Narang, and V. Chamola, “Reviewnet: A fast and resource optimized network for enabling safe autonomous driving in hazy weather conditions,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4256–4266, 2021.

[14] I. Bilik, O. Longman, S. Villeval, and J. Tabrikian, “The rise of radar for autonomous vehicles: Signal processing solutions and future research directions,” *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 20–31, 2019.

[15] W. Liu, M. Hua, Z. Deng, Z. Meng, Y. Huang, C. Hu, S. Song, L. Gao, C. Liu, B. Shuai, A. Khajepour, L. Xiong, and X. Xia, “A systematic survey of control techniques and applications in connected and automated vehicles,” 2023.

[16] L. Xiong, X. Xia, Y. Lu, W. Liu, L. Gao, S. Song, and Z. Yu, “Imu-based automated vehicle body sideslip angle and attitude estimation aided by gnss using parallel adaptive kalman filters,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 10 668–10 680, 2020.

[17] X. Xia, E. Hashemi, L. Xiong, and A. Khajepour, “Autonomous vehicle kinematics and dynamics synthesis for sideslip angle estimation based on consensus kalman filter,” *IEEE Transactions on Control Systems Technology*, vol. 31, no. 1, pp. 179–192, 2023.

[18] X. Xia, E. Hashemi, L. Xiong, A. Khajepour, and N. Xu, “Autonomous vehicles sideslip angle estimation: Single antenna gnss/imu fusion with observability analysis,” *IEEE Internet of Things Journal*, vol. 8, no. 19, pp. 14 845–14 859, 2021.

[19] W. Liu, X. Xia, L. Xiong, Y. Lu, L. Gao, and Z. Yu, “Automated vehicle sideslip angle estimation considering signal measurement characteristic,” *IEEE Sensors Journal*, vol. 21, no. 19, pp. 21 675–21 687, 2021.

[20] L. Liu, S. Lu, R. Zhong, B. Wu, Y. Yao, Q. Zhang, and W. Shi, “Computing systems for autonomous driving: State of the art and challenges,” *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6469–6486, 2021.

[21] S.-C. Lin, Y. Zhang, C.-H. Hsu, M. Skach, M. E. Haque, L. Tang, and J. Mars, “The architectural implications of autonomous driving: Constraints and acceleration,” *SIGPLAN Not.*, vol. 53, no. 2, p. 751–766, mar 2018. [Online]. Available: https://doi.org/10.1145/3296957.3173191

[22] L. Chen, Y. Li, C. Huang, Y. Xing, D. Tian, L. Li, Z. Hu, S. Teng, C. Lv, J. Wang, D. Cao, N. Zheng, and F.-Y. Wang, “Milestones in autonomous driving and intelligent vehicles—part i: Control, computing system design, communication, hd map, testing, and human behaviors,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 9, pp. 5831–5847, 2023.

[23] J. Chen, S. E. Li, and M. Tomizuka, “Interpretable end-to-end urban autonomous driving with latent deep reinforcement learning,” *CoRR*, vol. abs/2001.08726, 2020. [Online]. Available: https://arxiv.org/abs/2001.08726

[24] R. T. McAllister, Y. Gal, A. Kendall, M. van der Wilk, A. Shah, R. Cipolla, and A. Weller, “Concrete problems for autonomous vehicle safety: Advantages of bayesian deep learning,” in *International Joint Conference on Artificial Intelligence*, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:5631404

[25] D. Omeiza, H. Webb, M. Jirotka, and L. Kunze, “Explanations in autonomous driving: A survey,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 10 142–10 162, 2022.

[26] Éloi Zablocki, H. Ben-Younes, P. Pérez, and M. Cord, “Explainability of deep vision-based autonomous driving systems: Review and challenges,” 2022.

[27] S. Atakishiyev, M. Salameh, H. Yao, and R. Goebel, “Explainable artificial intelligence for autonomous driving: A comprehensive overview and field guide for future research directions,” 2023.

[28] H. Meinel and J. Dickman, “Automotive radar: from its origin to future directions,” *Microwave Journal*, vol. 56, no. 9, pp. 24–40, 2013.

[29] H. H. Meinel, “Evolving automotive radar — from the very beginnings into the future,” in *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*, 2014, pp. 3107–3114.

[30] I. Bilik, O. Longman, S. Villeval, and J. Tabrikian, "The rise of radar for autonomous vehicles: Signal processing solutions and future research directions," *IEEE signal processing Magazine*, vol. 36, no. 5, pp. 20–31, 2019.

[31] I. Bilik, O. Bialer, S. Villeval, H. Sharifi, K. Kona, M. Pan, D. Persechini, M. Musni, and K. Geary, "Automotive mimo radar for urban environments," in *2016 IEEE Radar Conference (RadarConf)*, 2016, pp. 1–6.

[32] M. Levy-Israel, I. Bilik, and J. Tabrikian, "Misspecified cramér-rao bound for multipath model in MIMO radar," in *2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2022, pp. 81–85.

[33] O. Longman, S. Villeval, and I. Bilik, "Multipath ghost targets mitigation in automotive environments," in *2021 IEEE Radar Conference (RadarConf21)*, 2021.

[34] L. Li and J. L. Krolik, "Simultaneous target and multipath positioning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 1, pp. 153–165, 2014.

[35] L. Elster, M. F. Holder, and M. Rapp, "A dataset for radar scattering characteristics of vehicles under real-world driving conditions: Major findings for sensor simulation," *IEEE Sensors Journal*, vol. 23, no. 5, pp. 4873–4882, 2023.

[36] Y. Li and X. Shang, "Multipath ghost target identification for automotive MIMO radar," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*, 2022.

[37] E. Tsalolikhin, I. Bilik, and N. Blaunstein, "A single-base-station localization approach using a statistical model of the NLOS propagation conditions in urban terrain," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 1124–1137, 2011.

[38] M. N. De Sousa and R. S. Thomä, "Mobile station localization emitter in urban NLOS using multipath ray tracing fingerprints and machine learning," in *2018 8th International Conference on Localization and GNSS (ICL-GNSS)*, 2018.

[39] J. Li, I.-T. Lu, and J. S. Lu, "Hybrid fingerprinting and ray extension localization in NLOS regions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 23 503–23 516, 2022.

[40] H. Naseri and V. Koivunen, "Cooperative simultaneous localization and mapping by exploiting multipath propagation," *IEEE Transactions on Signal Processing*, vol. 65, no. 1, pp. 200–211, 2017.

[41] D. Brodeski, I. Bilik, and R. Giryes, "Deep radar detector," in *2019 IEEE Radar Conference (RadarConf)*, 2019.

[42] F. Yavuz and M. Kalfa·, "Radar target detection via deep learning," in *2020 28th Signal Processing and Communications Applications Conference (SIU)*, 2020.

[43] C. x. Wang, Z. c. Chen, X. Chen, X. Tang, and F. t. Liang, "Detection of mmw radar target based on doppler characteristics and deep learning," in *2021 IEEE International Conference on Artificial Intelligence and Industrial Design (AIID)*, 2021, pp. 266–271.

[44] Y. Cheng, J. Su, H. Chen, and Y. Liu, "A new automotive radar 4d point clouds detector by using deep learning," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 8398–8402.

[45] J. F. Tilly, F. Weishaupt, O. Schumann, J. Dickmann, and G. Wanielik, "Road user detection on polarimetric pre-cfar radar data level," *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3558–3565, 2023.

[46] W. Kim, H. Cho, J. Kim, B. Kim, and S. Lee, "Target classification using combined yolo-svm in high-resolution automotive fmcw radar," in *2020 IEEE Radar Conference (RadarConf20)*, 2020.

[47] K. Patel, K. Rambach, T. Visentin, D. Rusev, M. Pfeiffer, and B. Yang, "Deep learning-based object classification on automotive radar spectra," in *2019 IEEE Radar Conference (RadarConf)*, 2019.

[48] M. Ulrich, C. Gläser, and F. Timm, "Deepreflecs: Deep learning for automotive object classification with radar reflections," in *2021 IEEE Radar Conference (RadarConf21)*, 2021.

[49] S. K. Suresh, B. Puttaswamappa, K. Rajarathinam, and P. Kamat, "Machine learning approach for obstacle classification in automotive radar," in *2022 1st International Conference on Informatics (ICI)*, 2022, pp. 86–91.

[50] A. Dubey, A. Santra, J. Fuchs, M. Lübke, R. Weigel, and F. Lurz, “A bayesian framework for integrated deep metric learning and tracking of vulnerable road users using automotive radars,” *IEEE Access*, vol. 9, pp. 68 758–68 777, 2021.

[51] M. S. Seyfioğlu, A. M. Özbayoğlu, and S. Z. Gürbüz, “Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1709–1723, 2018.

[52] N. Rojhani, M. Passafiume, M. Sadeghibakhi, G. Collodi, and A. Cidronali, “Model-based data augmentation applied to deep learning networks for classification of micro-doppler signatures using fmcw radar,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 71, no. 5, pp. 2222–2236, 2023.

[53] H.-W. Cho, S. Choi, Y.-R. Cho, and J. Kim, “Complex-valued channel attention and application in ego-velocity estimation with automotive radar,” *IEEE Access*, vol. 9, pp. 17 717–17 727, 2021.

[54] J. Mun, S. Ha, and J. Lee, “Automotive radar signal interference mitigation using rnn with self attention,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 3802–3806.

[55] M. Rameez, S. Javadi, M. Dahl, and M. I. Pettersson, “Signal reconstruction using bi-lstm for automotive radar interference mitigation,” in *2021 18th European Radar Conference (EuRAD)*, 2022, pp. 74–77.

[56] C. Jiang, Z. Zhou, and B. Yang, “Unsupervised deep interference mitigation for automotive radar,” in *2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2022, pp. 296–300.

[57] N.-C. Ristea, A. Anghel, and R. T. Ionescu, “Estimating the magnitude and phase of automotive radar signals under multiple interference sources with fully convolutional networks,” *IEEE Access*, vol. 9, pp. 153 491–153 507, 2021.

[58] J. Rock, W. Roth, M. Toth, P. Meissner, and F. Pernkopf, “Resource-efficient deep neural networks for automotive radar interference mitigation,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 927–940, 2021.

[59] S. Chen, J. Taghia, U. Kühnau, N. Pohl, and R. Martin, “A two-stage dnn model with mask-gated convolution for automotive radar interference detection and mitigation,” *IEEE Sensors Journal*, vol. 22, no. 12, pp. 12 017–12 027, 2022.

[60] J. Lee, D. Niyato, Y. L. Guan, and D. I. Kim, “Learning to schedule joint radar-communication with deep multi-agent reinforcement learning,” *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 406–422, 2022.

[61] S. Chen, M. Klemp, J. Taghia, U. Kühnau, N. Pohl, and R. Martin, “Improved target detection through dnn-based multi-channel interference mitigation in automotive radar,” *IEEE Transactions on Radar Systems*, 2023.

[62] M. M. Adnan, M. S. M. Rahim, A. Rehman, Z. Mehmood, T. Saba, and R. A. Naqvi, “Automatic image annotation based on deep learning models: A systematic review and future challenges,” *IEEE Access*, vol. 9, pp. 50 253–50 264, 2021.

[63] I. E. de Souza, C. Lessio Cazarin, M. R. Veronez, L. Gonzaga, and A. Xavier Falcão, “User-guided data expansion modeling to train deep neural networks with little supervision,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, 2022.

[64] S. Gannamaneni, S. Houben, and M. Akila, “Semantic concept testing in autonomous driving by extraction of object-level annotations from carla,” in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 2021, pp. 1006–1014.

[65] M. Dimitrievski, I. Shopovska, D. Van Hamme, P. Veelaert, and W. Philips, “Automatic labeling of vulnerable road users in multi-sensor data,” in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021, pp. 2623–2630.

[66] Z. Ying, C. Chao, W. Lei, Z. Shuai, and Z. XiangLei, “A data annotation and recognition method based on zero statistical hypothesis test and multi variable binary classification theory,” in *2022 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, 2022, pp. 676–680.

[67] S. Roychowdhury and L. S. Muppirisetty, “Fast proposals for image and video annotation using modified echo state

networks," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2018, pp. 1225–1230.

[68] S. Swetha, A. Mishra, G. M. Hegde, and C. Jawahar, "Efficient object annotation for surveillance and automotive applications," in *2016 IEEE Winter Applications of Computer Vision Workshops (WACVW)*, 2016.

[69] F. Schalling, S. Ljungberg, and N. Mohan, "Benchmarking lidar sensors for development and evaluation of automotive perception," in *2019 4th International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE)*, 2019.

[70] O. Schumann, M. Hahn, N. Scheiner, F. Weishaupt, J. F. Tilly, J. Dickmann, and C. Wöhler, "Radarscenes: A real-world radar point cloud data set for automotive applications," in *2021 IEEE 24th International Conference on Information Fusion (FUSION)*, 2021.

[71] R. Pfeffer, K. Bredow, and E. Sax, "Trade-off analysis using synthetic training data for neural networks in the automotive development process," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 4115–4120.

[72] R. van Kempen, B. Lampe, L. Reiher, T. Woopen, T. Beemelmanns, and L. Eckstein, "Data-driven occupancy grid mapping using synthetic and real-world data," in *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, 2022.

[73] N. Scheiner, N. Appenrodt, J. Dickmann, and B. Sick, "Automated ground truth estimation of vulnerable road users in automotive radar data using gnss," in *2019 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, 2019.

[74] A. Palffy, E. Pool, S. Baratam, J. F. P. Kooij, and D. M. Gavrila, "Multi-class road user detection with 3+1d radar in the view-of-delft dataset," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4961–4968, 2022.

[75] D. Scharf, S. Saadani, M. Schneider, and A. Höß, "A semi-automated multi-sensor data labeling process for deep learning in automotive radar environment," in *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, 2022.

[76] F. Kraus, N. Scheiner, W. Ritter, and K. Dietmayer, "The radar ghost dataset–an evaluation of ghost objects in automotive radar data," in *in Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 8570–8577.

[77] M. Chamseddine, J. Rambach, D. Stricker, and O. Wasenmuller, "Ghost target detection in 3D radar data using point cloud based deep neural network," in *in Proc. 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 10 398–10 403.

[78] L. Stäcker, S. Mishra, P. Heidenreich, J. Rambach, and D. Stricker, "Rc-bevfusion: A plug-in module for radar-camera bird's eye view feature fusion," 2023.

[79] Y. Zhou, L. Liu, H. Zhao, M. López-Benítez, L. Yu, and Y. Yue, "Towards deep radar perception for autonomous driving: Datasets, methods, and challenges," *Sensors*, vol. 22, no. 11, p. 4208, 2022.

[80] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," 2020.

[81] E. Pronovost, K. Wang, and N. Roy, "Generating driving scenes with diffusion," 2023.

[82] S. Chen, P. Sun, Y. Song, and P. Luo, "Diffusiondet: Diffusion model for object detection," 2023.

[83] J. Sohl-Dickstein, E. A. Weiss, N. Maheswaranathan, and S. Ganguli, "Deep unsupervised learning using nonequilibrium thermodynamics," 2015.

[84] M. Najibi, J. Ji, Y. Zhou, C. R. Qi, X. Yan, S. Ettinger, and D. Anguelov, "Unsupervised 3d perception with 2d vision-language distillation for autonomous driving," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2023, pp. 8602–8612.

[85] X. Ding, J. Han, H. Xu, W. Zhang, and X. Li, "Hilm-d: Towards high-resolution understanding in multimodal large language models for autonomous driving," 2023.

[86] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," 2014.

[87] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," 2019.

[88] L. Yao, R. Huang, L. Hou, G. Lu, M. Niu, H. Xu, X. Liang, Z. Li, X. Jiang, and C. Xu, "Filip: Fine-grained interactive language-image pre-training," 2021.

[89] J. Zhang, J. Huang, S. Jin, and S. Lu, "Vision-language models for vision tasks: A survey," 2023.

[90] C. Schuhmann, R. Beaumont, R. Vencu, C. Gordon, R. Wightman, M. Cherti, T. Coombes, A. Katta, C. Mullis, M. Wortsman, P. Schramowski, S. Kundurthy, K. Crowson, L. Schmidt, R. Kaczmarczyk, and J. Jitsev, "Laion-5b: An open large-scale dataset for training next generation image-text models," 2022.

[91] Y. Du, K. Konyushkova, M. Denil, A. Raju, J. Landon, F. Hill, N. de Freitas, and S. Cabi, "Vision-language models as success detectors," 2023.

[92] Z. Xu, C. Tang, and M. Tomizuka, "Zero-shot deep reinforcement learning driving policy transfer for autonomous vehicles based on robust control," pp. 2865–2871, 2018.

[93] E. Miller, N. Matsakis, and P. Viola, "Learning from one example through shared densities on transforms," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, vol. 1, 2000, pp. 464–471 vol.1.

[94] G. R. Koch, "Siamese neural networks for one-shot image recognition," 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:13874643

[95] M. Rezaei and M. Shahidi, "Zero-shot learning and its applications from autonomous vehicles to COVID-19 diagnosis: A review," *Intelligence-Based Medicine*, vol. 3-4, p. 100005, dec 2020.

[96] P. Zhou, J. Zhu, Y. Wang, Y. Lu, Z. Wei, H. Shi, Y. Ding, Y. Gao, Q. Huang, Y. Shi, A. Alhilal, L.-H. Lee, T. Braud, P. Hui, and L. Wang, "Vetaverse: A survey on the intersection of metaverse, vehicles, and transportation systems," 2023.

[97] H. Zhang, S. Mao, D. Niyato, and Z. Han, "Location-dependent augmented reality services in wireless edge-enabled metaverse systems," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 171–183, 2023.

[98] M. Xu, W. C. Ng, W. Y. B. Lim, J. Kang, Z. Xiong, D. Niyato, Q. Yang, X. S. Shen, and C. Miao, "A full dive into realizing the edge-enabled metaverse: Visions, enabling technologies,and challenges," 2022.

[99] Y. Hui, X. Ma, Z. Su, N. Cheng, Z. Yin, T. H. Luan, and Y. Chen, "Collaboration as a service: Digital-twin-enabled collaborative and distributed autonomous driving," *IEEE Internet of Things Journal*, vol. 9, no. 19, pp. 18 607–18 619, 2022.

[100] A. Hu, L. Russell, H. Yeo, Z. Murez, G. Fedoseev, A. Kendall, J. Shotton, and G. Corrado, "Gaia-1: A generative world model for autonomous driving," 2023.

[101] M. Xu, D. Niyato, J. Chen, H. Zhang, J. Kang, Z. Xiong, S. Mao, and Z. Han, "Generative ai-empowered simulation for autonomous driving in vehicular mixed reality metaverses," 2023.

[102] S. Cakir, M. Gauß, K. Häppeler, Y. Ounajjar, F. Heinle, and R. Marchthaler, "Semantic segmentation for autonomous driving: Model evaluation, dataset generation, perspective comparison, and real-time capability," 2022.

[103] S. Glass, I. Mahgoub, and M. Rathod, "Leveraging manet-based cooperative cache discovery techniques in vanets: A survey and analysis," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2640–2661, 2017.

[104] D. Jia, K. Lu, J. Wang, X. Zhang, and X. Shen, "A survey on platoon-based vehicular cyber-physical systems," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 263–284, 2016.

[105] A. Boualouache and T. Engel, “A survey on machine learning-based misbehavior detection systems for 5g and beyond vehicular networks,” *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1128–1172, 2023.

[106] J. Petit, F. Schaub, M. Feiri, and F. Kargl, “Pseudonym schemes in vehicular networks: A survey,” *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 228–255, 2015.

[107] H. Khelifi, S. Luo, B. Nour, H. Moungla, Y. Faheem, R. Hussain, and A. Ksentini, “Named data networking in vehicular ad hoc networks: State-of-the-art and challenges,” *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 320–351, 2020.

[108] S. Dietzel, J. Petit, F. Kargl, and B. Scheuermann, “In-network aggregation for vehicular ad hoc networks,” *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1909–1932, 2014.

[109] J. Wang, Z. Yan, H. Wang, T. Li, and W. Pedrycz, “A survey on trust models in heterogeneous networks,” *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2127–2162, 2022.

[110] K. Abboud, H. A. Omar, and W. Zhuang, “Interworking of dsrc and cellular network technologies for v2x communications: A survey,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 9457–9470, 2016.