

Towards Real-time Learning for Edge-Cloud Continuum with Vehicular Computing

Ella Peltonen[†], Arun Sojan[‡], Tero Päivärinta[‡]

[†]*Center for Ubiquitous Computing, University of Oulu, Finland*

[‡]*M3S Research Unit, University of Oulu, Finland*

firstname.lastname@oulu.fi

Abstract—Sensor-driven IoT systems are well-known for their capacity to accelerate massive amounts of data in a comparatively short period of time. To have any use, the information delivery and decision making based on the data require efficient learning models together with dynamically deployed computing and network resources. The current cloud and high-performance computing infrastructures, as well as modern edge computing systems especially in the 5G and beyond networks, can be addressed to resolve these challenges. However, there are several application areas especially in vehicular and urban computing, where just harnessing more computational power does not solve computational and real-time requirements of the modern sensing systems that operate in mobile and context-dependent environments. For now, the mathematical challenges of distributed computing and real-time learning algorithms have not been profoundly addressed in the context of the IoT and real-world sensing applications. Data-driven systems also require giving full attention to information delivery, data management, data cleaning, and sensor fusion technologies that need to be equally distributed and real-time competent as the learning algorithms themselves. New software-defined computing and networking approaches and architectures are required to orchestrate the numerous connected resources dynamically, controllably, and securely along with the evolving needs. The key challenge here is to uniform collaboration between different aspects of the system, from data processing and delivery to the algorithms and learning models, not forgetting the computational capacity and networking capabilities, all this in real-time with real-world applications.

Index Terms—Edge-Cloud Continuum, Internet of Things, Vehicular Computing, Machine Learning, Artificial Intelligence, Automotive Software, Software-Defined Vehicle, Software-Defined Network, Smart Traffic.

I. INTRODUCTION

Today, driving is safer than ever. Modern cars implement technologies for automatic braking, Cooperative Adaptive Cruise Control (CACC), prevention of unwanted lane crossing, distance keeping, and so on, to supply driver's own cognition and prevent accidents. However, 239 people died and 956 people were wounded in traffic-related accidents in 2018 alone in Finland. Based on the Commission reports, 25 100 people lost their lives on EU roads in 2018 and about 135 000 were seriously injured. The EU has reaffirmed an ambitious long-term goal, to move close to zero deaths in traffic by 2050 i.e. "Vision Zero" [1]. One of the main steps to achieve this goal is safer vehicles, specifically, connected, and automated vehicles interacting with other road users and with the digital and physical road infrastructure [2].

Connected vehicles and vehicular edge computing is seen as a potential solution for providing services and applications for safe driving [3]. With increased networking and computing capabilities, connected vehicles and vehicular edge networks can perform challenging inference and learning task to support driver's cognition and automate the driving scenario. Such intelligent systems demand training data, which can be provided by in-vehicle sensors and external databases as shown in Figure 1. However, how to combine this information in a challenging real-time and mobile environment with modern software-defined networks (SDNs) and computing resources is still an open question.

In the context of vehicular sensor analytics, machine learning (ML) and artificial intelligence (AI) solutions have previously been used to improve lane [4] and road potholes [5] recognition. These existing systems usually utilise only in-vehicle sensors [6], such as camera, LiDARs, radars, and speed meters [7], [8]. Driver's behaviour is previously detected in the context of smartphone usage during driving [9] and drunk driving [10], however, underlining that generally human driver's perception and reasoning still maintain an advantage compared to fully automatically controlled vehicles [11]. To understand the whole picture of the driving performance, both in-vehicle sensors and driver's behavioural patterns have to be studied. Literally, local real-time computing can ease driver's senses to see "around the corner" and detect potentially hazardous situations. In our vision, we envisage simultaneous sensing of the driver, vehicle, and driving environment, for which multi-sourced sensor analysis and ML/AI approaches are integrated into the vehicular edge-cloud continuum.

Vehicular sensors can produce a large amount of data in a relatively short time [12] by which the current paradigm is to utilise fast data transmission technologies [13], such as 5G networks [14], [15] and edge computing services [16], [17], to replace or supplement traditional cloud computing services. The vehicular environment, consisting of the car's own computers, driver's smart devices, and road-side contact points, holds a greatly potential computational capacity that is, however, still unrecognised as a part of vehicular edge computing [18], [16]. Indeed, several ML/AI models can be brought to the automotive software system itself [19]. The main question here is the management of the dynamic orchestration of the vehicular edge-cloud continuum, or in layman terms, what to compute, when, and where.

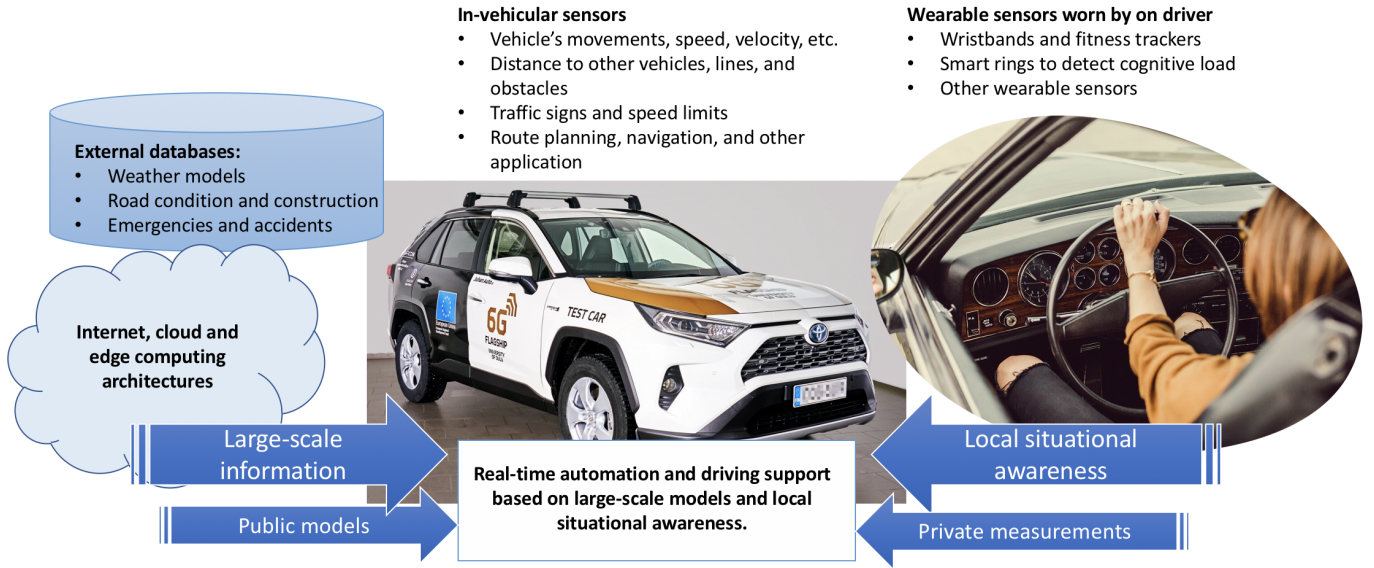


Fig. 1. An overview of the multiple data sources in a vehicular computing environment; public data from internet have to be fused with private sensor-driven data expanding the situational awareness of the system.

A novel research field of vehicular edge computing has emerged only recently [3], [6]. With networking capabilities in 5G and beyond, connected vehicles themselves can be seen as a part of such an edge system and, on a larger scale, computational units for the vehicular edge-cloud continuum. However, in-vehicular software systems and specialised hardware components, such as GPU-capable AGX boards, are relatively new on market and their ML/AI capabilities are not fully evaluated and tested in practice. Specialised hardware solutions can, on the other hand, be expensive but their role can be seen to become more integral in automotive development. The question remains where to compute which parts of the inference and decision making logic. There are several open questions on how to enable a fully operational pipeline from collecting the multi-sourced sensing data, learning the ML/AI models, and utilising the models for feedback and decision making.

In this paper, we aim to conceptualise the problems emerging in software and system developments where challenging data management, machine learning, and analytics solutions need to fit together with very dynamic and changing software constellations. To summarise, we present the following contributions:

- We outline the challenges of bringing sensing-based ML/AI model training and inference in the vehicular edge itself in a distributed fashion (contrasting to the current cloud-centric paradigm).
- We present a reference architecture for the ML/AI pipelines for the edge-cloud continuum while maintaining its manageability through the software-defined approach.

II. VISION STATEMENT AND RESEARCH AGENDA

Bringing sensing-based ML/AI solution into the vehicular computing environment is largely understudied what comes

to the challenges of real-time requirements [20], [13], high mobility of the clients [21], [22], and a need for task allocation [16], [23]. Vehicular edge computing is an especially important development area because it enables various critical aspect for safer driving [24], [3]: Low latency enables real-time feedback and various safety applications. In remote areas and while on the move, connectivity lost to the cloud prevents decision and information sharing, whereas, on the edge, local resources are effectively harnessed for more resilient and fault-tolerant computing. The solutions to these challenges can further be generalised for other real-world use cases involving mobility or real-time demands, such as smart hospitals, manufacturing, and Industry 4.0. Based on this rationale, we emphasise the following open challenges to be solved before the ML/AI pipelines can be fully integrated into the vehicular edge-cloud continuum:

Resource-efficient ML/AI and data analytics need to meet the high mobility requirements of the vehicular environment. The current state of the art for the sensor-driven ML/AI solutions involves collecting the data from the clients and sending it to a remote location, usually a cloud computing service, for analysis, inference and prediction [25], [26], [27]. This is due to the easy-to-use and common cloud services [28] and concern of the battery life and energy efficiency of the client devices [29]. However, energy-aware distributed algorithms can take advantage local computational power available on the edge [24] and vehicles as a part of it. Indeed, there are even energy benefits to gain from localised data processing [30]. With the decreased need for data transmission, both energy and time benefits can be gained. However, the orchestrating system needs to dynamically consider existing capabilities and resources in the system, including the vehicle itself and available edge

services. Thus, intelligent vehicular systems need to become context-aware.

Context-aware sensing solutions are needed to meet the real-time requirements. Large-scale sensor data analytics and ML/AI solutions currently rely on novel algorithmic innovations, especially federated learning [31], [32] and other highly popular technologies [33], [34], [35]. However, with the real-time requirements, more transparent results can be achieved by relatively traditional statistical analysis tools, such as regression and anomaly detection algorithms [36], [19]. For this, the full understanding of the vehicle's operating context, requirements, resources, and limitations, need to be understood. Based on the context-aware decision making, the most lightweight solution can be used to provide as good or – especially in terms of computation time and resource-efficiency – even better results in the rapidly changing contextual conditions [37]. Withing the dynamic intelligent solutions and capabilities, there is an increasing need for ML/AI-fluent DevOps practices and software pipelines that can consider different types of ML/AI models, data processing steps, and data management.

There is a need for a full development pipeline for edge-native and in-vehicle ML/AI, aiming towards an intelligent vehicular edge-cloud continuum. Light-weight ensemble learning methods can be used to create a set of models to be chosen from, fitting to the dynamic nature of the mobile and real-time vehicle-edge-cloud environment [38], [39], [40]. However, automotive software and hardware solutions still require work towards a fully integrated development pipeline. Here, different parts of data processing need to meet in an effective operational loop: 1) data collection from various sources (including in-vehicle sensors and Internet-based databases, such as weather and road condition data), 2) data cleaning and preprocessing steps integral for real-life data, 3) distributed and online ML/AI model building on the vehicular edge network, and 4) real-time inference and decision making without centralised support in the cloud. In addition, security and privacy aspects need to be addressed [41].

Software-defined approach to configuration, orchestration, and maintenance of the required sensor, data processing, storage, and network resources and the resulting dynamic and complex systems of interacting vehicles (with their sub-components), edge nodes, and cloud services is required to make the vehicle-edge-cloud continuum possible in the first place. Numerous developments on software-defined solutions and reference architectures for vehicular and edge-cloud computing have been suggested recently. The concept of the software-defined vehicle is required to manage the numerous electronic control units, sensors and their connections through in-vehicle networks alone [42], [43]. Furthermore, the software-defined networking approach emerges as a crucial pre-requisite for managing vehicular ad-hoc networks of connected vehicles necessary for the evolving smart traffic and transport systems [44]. The offloading and external data services then require software-defined, multi-access edge and cloud computing solutions [45], [41] to be dynamically, scalably, and securely orchestrated with the vehicular ad-hoc networks.

III. TOWARDS A REFERENCE ARCHITECTURE

For considering the ML/AI pipelines for the edge-cloud continuum while maintaining its manageability through the software-defined approach, especially from the automotive perspective, we propose the following statements as design principles for the possible reference architecture:

- 1) Extend the software on automotive so that it becomes possible to run ML/AI tasks considering various data sources for learning processes and add capabilities to offload ML/AI tasks to MEC or cloud-based on the task complexity, on-board capability, time sensitivity and availability of the MEC;
- 2) Extend parallel and distributed ML/AI methods to bring together driving patterns, in-vehicle sensors, and external databases in a real-time manner;
- 3) Extend evaluation practices for the feasibility of sensor-driven ML/AI for vehicular computing to improve driving safety and human driver's context-awareness.
- 4) To achieve the three above-mentioned design principles, adopt the software-defined approach to orchestrate the computing elements within the vehicle, ad-hoc vehicular networks, and the networking and resource management of offloaded edge and cloud services.

As shown in Figure 2, several parallel data-intensive tasks need to be included in the potential reference architecture. These are:

- 1) In-vehicular sensor and human interaction data gathering in a real-time manner from data streams, as well as receiving the relevant data items in branches from the MEC and cloud back-end.
- 2) Data processing tasks individually for different data sources, consisting of data cleaning, preprocessing, and validation pipelines.
- 3) In-sensor, local (CPU or GPU), or edge-supported ML/AI model training and validation with preprocessed data. In addition, model up-keeping and life cycle need to be considered to remove unnecessary information.
- 4) Fusing information coming from different data streams and ML/AI models.
- 5) Inference and decision making based on the provided knowledge from the model fusion, including all the available information from both external and internal data sources.
- 6) Offloading to other layers based on the analysis of complexity, time sensitivity, layer capability identification and availability.

As seen in Figure 2, edge and cloud parts of the operating system can still largely be utilised to process large-scale datasets, such as weather and road condition information. In addition, layers have the capability to selectively offload tasks to the other layers if necessary. Common to these data sources is that they benefit collaborative and pervasive data collection techniques including various sensor networks (such as roadside cameras) and existing databases (such as weather models). However, local in-vehicle sensors (including driver-adjusted

vehicle computing resources and networks, dynamic ad-hoc vehicular networks, and related edge-cloud resources. The development is foreseen to enable real-time, mission-critical, context-aware, and efficient applications of the vehicular edge-cloud computing continuum.

ACKNOWLEDGEMENTS

This research is supported by the Academy of Finland 6Genesis Flagship (grant number 318927) and SMAD Project.

REFERENCES

- [1] E. Commission, "White Paper "Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system", 2011, cOM(2011) 144 final.
- [2] —, "Staff Working Paper EU Road Safety Policy Framework 2021-2030 – Next steps towards "Vision Zero", 2019, sWD(2019) 283 final.
- [3] A. Alhilal, T. Braud, and P. Hui, "Distributed vehicular computing at the dawn of 5g: a survey," *arXiv:2001.07077*, 2020.
- [4] Q. Li, L. Chen, M. Li, S. L. Shaw, and A. Nüchter, "A sensor-fusion drivable-region and lane-detection system for autonomous vehicle navigation in challenging road scenarios," *IEEE T. on Vehicular Technology*, vol. 63, no. 2, pp. 540–555, 2014.
- [5] A. Ghose, P. Biswas, C. Bhaumik, M. Sharma, A. Pal, and A. Jha, "Road condition monitoring and alert application," in *IEEE Int. Conf. on Pervasive Computing and Communications Workshops*. Lugano, Switzerland: IEEE, 2012, pp. 489–491.
- [6] S. Baidya, Y. Ku, H. Zhao, J. Zhao, and S. Dey, "Vehicular and edge computing for emerging connected and autonomous vehicle applications," in *Design Automation Conf. (DAC)*, 2020.
- [7] M. Munz, M. Mahlisch, and K. Dietmayer, "Generic centralized multi sensor data fusion based on probabilistic sensor and environment models for driver assistance systems," *IEEE Intelligent Transportation Systems*, vol. 2, no. 1, 2010.
- [8] F. Garcia, D. Martin, A. De La Escalera, and J. M. Armingol, "Sensor fusion methodology for vehicle detection," *IEEE Int Trans Sys*, vol. 9, no. 1, 2017.
- [9] Y. Wang, J. Yang, H. Liu, Y. Chen, M. Gruteser, and R. P. Martin, "Sensing vehicle dynamics for determining driver phone use," in *Int. conf. on mobile systems, applications, and services*, 2013, pp. 41–54.
- [10] J. Ljungblad, B. Hök, A. Allalou, and H. Pettersson, "Passive in-vehicle driver breath alcohol detection using advanced sensor signal acquisition and fusion," *Traffic injury prevention*, vol. 18, 2017.
- [11] B. Schoettle, "Sensor fusion: A comparison of sensing capabilities of human drivers and highly automated vehicles," *University of Michigan*, 2017.
- [12] N. C. et al., "Big data driven vehicular networks," *IEEE Network*, vol. 32, no. 6, pp. 160–167, 2018.
- [13] Y. Lai, H. Lin, F. Yang, and T. Wang, "Efficient data request answering in vehicular ad-hoc networks based on fog nodes and filters," *Future Generation Computer Systems*, vol. 93, pp. 130–142, 2019.
- [14] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally, "5g for vehicular communications," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 111–117, 2018.
- [15] X. Huang, R. Yu, J. Kang, Y. He, and Y. Zhang, "Exploring mobile edge computing for 5g-enabled software defined vehicular networks," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 55–63, 2017.
- [16] L. Liu, C. Chen, Q. Pei, S. Maharjan, and Y. Zhang, "Vehicular edge computing and networking: A survey," *Mobile Networks and Applications*, 2020.
- [17] G. Luo, Q. Yuan, H. Zhou, N. Cheng, Z. Liu, F. Yang, and X. S. Shen, "Cooperative vehicular content distribution in edge computing assisted 5g-vanet," *China Communications*, vol. 15, no. 7, pp. 1–17, 2018.
- [18] J. Kang, R. Yu, X. Huang, M. Wu, S. Maharjan, S. Xie, and Y. Zhang, "Blockchain for secure and efficient data sharing in vehicular edge computing and networks," *IEEE Internet of Things Journal*, vol. 6, no. 3, 2018.
- [19] G. Bresson, M.-C. Rahal, D. Gruyer, M. Revilloud, and Z. Alsayed, "A cooperative fusion architecture for robust localization: Application to autonomous driving," in *Int Conf on Intelligent Transportation systems (ITSC)*. IEEE, 2016.
- [20] T. S. Darwish and K. A. Bakar, "Fog based intelligent transportation big data analytics in the internet of vehicles environment: motivations, architecture, challenges, and critical issues," *IEEE Access*, vol. 6, 2018.
- [21] T. D. Nguyen, T.-D. Nguyen, V. D. Nguyen, X.-Q. Pham, E.-N. Huh *et al.*, "Cost-effective resource sharing in an internet of vehicles-employed mobile edge computing environment," *Symmetry*, vol. 10, no. 11, 2018.
- [22] Y. Dai, D. Xu, S. Maharjan, and Y. Zhang, "Joint load balancing and offloading in vehicular edge computing and networks," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4377–4387, 2018.
- [23] Z. Zhou, P. Liu, J. Feng, Y. Zhang, S. Mumtaz, and J. Rodriguez, "Computation resource allocation and task assignment optimization in vehicular fog computing: A contract-matching approach," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, 2019.
- [24] J. Zhang and K. B. Letaief, "Mobile edge intelligence and computing for the internet of vehicles," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 246–261, 2019.
- [25] Y. Simmhan, S. Aman, A. Kumbhare, R. Liu, S. Stevens, Q. Zhou, and V. Prasanna, "Cloud-based software platform for big data analytics in smart grids," *Computing in Science & Engineering*, vol. 15, no. 4, p. 38, 2013.
- [26] M. M. Rathore, A. Ahmad, A. Paul, and S. Rho, "Urban planning and building smart cities based on the internet of things using big data analytics," *Computer Networks*, vol. 101, pp. 63–80, 2016.
- [27] M. Marjani, F. Nasaruddin, A. Gani, A. Karim, I. A. T. Hashem, A. Siddiq, and I. Yaqoob, "Big IoT data

- analytics: architecture, opportunities, and open research challenges,” *IEEE Access*, vol. 5, pp. 5247–5261, 2017.
- [28] A. Botta, W. De Donato, V. Persico, and A. Pescapé, “Integration of cloud computing and internet of things: a survey,” *Fut. gen. comp. sys.*, vol. 56, 2016.
- [29] K. Gai, M. Qiu, H. Zhao, L. Tao, and Z. Zong, “Dynamic energy-aware cloudlet-based mobile cloud computing model for green computing,” *J. of Network and Computer Applications*, vol. 59, pp. 46–54, 2016.
- [30] E. Peltonen, E. Lagerspetz, P. Nurmi, and S. Tarkoma, “Energy modeling of system settings: A crowdsourced approach,” in *IEEE Int. Conf. on Pervasive Computing and Communications*. IEEE, 2015, pp. 37–45.
- [31] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Konecny, S. Mazzocchi, H. B. McMahan, T. Van Overveldt, D. Petrou, D. Ramage, and J. Roselander, “Towards federated learning at scale: system design,” in *SysML Conf.*, Palo Alto, US, 2019.
- [32] J. Park, S. Samarakoon, M. Bennis, and M. Debbah, “Wireless network intelligence at the edge,” *arXiv:1812.02858*, 2018.
- [33] N. D. Lane, S. Bhattacharya, P. Georgiev, C. Forlivesi, and F. Kawsar, “An early resource characterization of deep learning on wearables, smartphones and internet-of-things devices,” in *Int. w. on internet of things towards applications*. ACM, 2015, pp. 7–12.
- [34] C. A. Ronao and S.-B. Cho, “Human activity recognition with smartphone sensors using deep learning neural networks,” *Expert systems with applications*, vol. 59, pp. 235–244, 2016.
- [35] Y. Liang, Z. Cai, J. Yu, Q. Han, and Y. Li, “Deep learning based inference of private information using embedded sensors in smart devices,” *IEEE Network*, vol. 32, no. 4, pp. 8–14, 2018.
- [36] P. J. Rousseeuw and M. Hubert, “Anomaly detection by robust statistics,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 2, 2018.
- [37] D. Ravi, C. Wong, B. Lo, and G.-Z. Yang, “A deep learning approach to on-node sensor data analytics for mobile or wearable devices,” *IEEE J. of biomedical and health informatics*, vol. 21, no. 1, pp. 56–64, 2016.
- [38] L. Lovén, T. Leppänen, E. Peltonen, J. Partala, E. Harjula, P. Poramberg, M. Ylianttila, and J. Riekk, “EdgeAI: A vision for distributed, edge-native artificial intelligence in future 6G networks,” in *The 1st 6G Wireless Summit*, Levi, Finland, 2019, pp. 1–2.
- [39] L. Lovén, E. Peltonen, T. Leppänen, E. Gilman, A. Pandya, S. Pirttikangas, and J. Riekk, “Towards Edison: An edge-native approach to distributed interpolation of environmental data,” in *Enabling Internet of Things Ecosystems through Edge Computing*, Spain, 2019.
- [40] L. Lovén, V. Karsisto, H. Järvinen, M. J. Sillanpää, T. Leppänen, E. Peltonen, S. Pirttikangas, and J. Riekk, “Mobile road weather sensor calibration by sensor fusion and linear mixed models,” *PloS one*, vol. 14, no. 2, 2019.
- [41] B. Ali, M. A. Gregory, and S. Li, “Multi-access edge computing architecture, data security and privacy: A review,” *IEEE Access*, vol. 9, pp. 18 706–18 721, 2021.
- [42] N. Cvijetic and T. Tomazin, “Developing a centralized compute architecture for autonomous vehicles,” *ATZ electronics worldwide*, vol. 16, pp. 10–15, 2021.
- [43] M. Haeberle, F. Heimgaertner, H. Loehr, N. Nayak, D. Grewe, S. Schildt, and M. Menth, “Softwarization of automotive e/e architectures: A software-defined networking approach,” in *IEEE Vehicular Networking Conf. (VNC)*. IEEE, 2020, pp. 1–8.
- [44] J. Bhatia, Y. Modi, S. Tanwar, and M. Bhavsar, “Software defined vehicular networks: A comprehensive review,” *Int. J. of Communication Systems*, vol. 32, pp. 1–22, 2019.
- [45] A. Wang, Z. Zha, Y. Guo, and S. Chen, “Software-defined networking enhanced edge computing: A network-centric survey,” *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1500–1519, 2019.
- [46] L. Chen, M. Torstensson, and C. Englund, “Federated learning to enable automotive collaborative ecosystem: opportunities and challenges,” in *Virtual ITS European Congress, Online, November 9-10, 2020*, 2020.
- [47] H. Zhang, J. Bosch, and H. H. Olsson, “Real-time end-to-end federated learning: An automotive case study,” *arXiv preprint arXiv:2103.11879*, 2021.