Ahmad Khalil, Tobias Meuser, Yassin Alkhalili, Antonio Fernandez Anta, Lukas Staecker, Ralf Steinmetz. *In Proceedings of the 8th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2022), pages 346-352*

Situational Collective Perception: Adaptive and Efficient Collective Perception in Future Vehicular Systems

Ahmad Khalil[1][a], Tobias Meuser[1][b], Yassin Alkhalili[1][c], Antonio Fernandez Anta[2][d], Lukas Staecker[3][e], and Ralf Steinmetz[1][f]

[1] *Multimedia Communications Lab, Technical University of Darmstadt, Darmstadt, Germany*

[2] *IMDEA Networks Institute, Madrid, Spain*

[3] *Stellantis, Opel Automobile GmbH, Rüsselsheim, Germany*

{*ahmad.khalil, tobias.meuser, yassin.alkhalili,ralf.steinmetz*}*@kom.tu-darmstadt.de, antonio.fernandez@imdea.org, lukas.staecker@external.stellantis.com*

Keywords: Collective Perception, Vehicular Networks, Intelligent Transportation Systems, V2X, Federated Learning

Abstract: **With the emerge of Vehicle-to-everything (V2X) communication, vehicles and other road users can perform Collective Perception (CP), whereby they exchange their individually detected environment to increase the collective awareness of the surrounding environment. To detect and classify the surrounding environmental objects, preprocessed sensor data (e.g., point-cloud data generated by a Lidar) in each vehicle is fed and classified by onboard Deep Neural Networks (DNNs). The main weakness of these DNNs is that they are commonly statically trained with context-agnostic data sets, limiting their adaptability to specific environments. This may eventually prevent the detection of objects, causing safety disasters. Inspired by the Federated Learning (FL) approach, in this work we tailor a collective perception architecture, introducing Situational Collective Perception (SCP) based on dynamically trained and situational DNNs, and enabling adaptive and efficient collective perception in future vehicular networks.**

1 INTRODUCTION

The number of vehicles on the roads is expected to reach 1.8 billion worldwide by 2035 (Löffler, 2021), therefore more crowded roads will continue to be a common phenomenon in the future. Crowded roads not only increase the probability of having accidents but also reduce the overall traffic efficiency. This highlights the necessity to develop and deploy more intelligent systems in the vehicles, to assist the vehicle's driver. Supporting intelligent systems help increasing traffic safety and efficiency, by making the vehicles' actions less dependent, or even fully independent, from the driver decisions.

However, when taking over, the intelligent systems in the vehicles need to be aware of their environment. By depending only on their onboard sensors, the vehicles have limited perception of their surrounding environment. Therefore, enabling communication to exchange perception data between vehicles and other road users is a pivotal aspect to ensure reaching the intended level of environmental awareness.

With the emerging Vehicle-to-everything (V2X) technology, road users can exchange data, which opens the horizons for a multitude of applications (Boban et al., 2018), like Cooperative Awareness (Sjoberg et al., 2017), Cooperative Maneuver (3GPP, 2016), Teleoperated Driving (3GPP, 2016), and Collective Perception (Shan et al., 2021; Fukatsu and Sakaguchi, 2019; Pilz et al., 2021; Shan et al., 2021; Barbieri et al., 2021).

In this work, we focus on some of today's issues of Collective Perception (CP) and introduce a possible solution for enhancing the vehicles' perception in specific environments.

Having different perception levels of the surrounding environment, vehicles and other road users can participate in the collective perception by exchanging their detected objects. This helps the road users to extend

[a] https://orcid.org/0000-0002-6059-7027
[b] https://orcid.org/0000-0002-2008-5932
[c] https://orcid.org/0000-0001-5722-1910
[d] https://orcid.org/0000-0001-6501-2377
[e] https://orcid.org/0000-0001-7537-606X
[f] https://orcid.org/0000-0002-6839-9359

Ahmad Khalil, Tobias Meuser, Yassin Alkhalili, Antonio Fernandez Anta, Lukas Staecker, Ralf Steinmetz. *In Proceedings of the 8th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2022), pages 346-352*

their perception range including the objects outside their Field of view (FOV) (e.g., objects hidden by a building). Nevertheless, there is a wide set of challenges within CP that needs to be considered. As CP is performed in three main steps, *sensing*, *communicating*, and *fusing* (Pilz et al., 2021), challenges can arise in each one of these steps. Moreover, having diverse levels of sensing technologies on the vehicles, the huge amount of transmitted data (Fukatsu and Sakaguchi, 2019), and the use of statically trained onboard Deep Neural Networks (DNNs) used for objects detection, are open challenges that have to be tackled by researchers.

In this work, we handle the issue of having onboard DNNs which are statically trained with context-agnostic data sets. These DNNs are statically trained and deployed on the vehicles, which could affect their detection performance in some certain circumstances (e.g., fog, heavy rain at night). Moreover, DNNs suffer from the high costs of performance improvement (NEIL et al., 2021), as the amount of data required for generalized DNNs scale exponentially with the required accuracy of these models.

Federated Learning (FL) has recently attracted considerable interest in the field of vehicular networks research, and more literature has been published in recent years showing its ability to train DNNs with a less amount of transmitted data (Otoum et al., 2020; Yu et al., 2020; Yuan et al., 2021; Boualouache and Engel, 2021; Zhao et al., 2021; Ding et al., 2022). Nonetheless, previous research has tended to focus on using FL to dynamically train generalized vehicles' onboard models and failed to make them context-aware.

To alleviate such issues, and inspired by FL concepts, the main contribution of this work is to introduce our forward-looking Situational Collective Perception (SCP), which enables dynamically-trained and situation-aware onboard DNNs. We tailor the CP architecture to enable training the situational DNNs with context-specific data, and deploy it on the vehicles to enhance their detection capability in specific situations.

The rest of the paper is organized as follows: In Section 2 we provide more insight into CP, its elements, key features, and its main challenges. Section 3 presents an overview of FL and highlights the recent works on applying FL concepts in vehicular networks. We introduce Situational Collective Perception (SCP) in Section 4, presenting its potential for enhancing the detection capability of the vehicles. In this section, we also highlight the possible future research directions and challenges of adopting SCP in vehicular networks. Finally, in Section 5, we conclude this work.

2 COLLECTIVE PERCEPTION

The term collective perception refers to one of the emerging technologies in Intelligent Transportation Systems (ITSs). Figure 1 illustrates the main components of collective perception. Each vehicle uses its onboard sensors (e.g., camera, lidar, radar) to gather data about the surrounding environment. The generated raw sensor data is preprocessed (e.g., size reduction, fusion) and fed to Deep Neural Networks (DNNs) for object detection and classification. The vehicle uses a list of the detected objects and merges it with their spatial information to build its local environment model. This local environment model is the basis of decision-making and planning systems.

However, the main weakness of this individual perception (without data exchange between vehicles) is that it highly depends on the onboard sensor technology (obstruction, detection range, precision), which restricts the perception range to each vehicle's Field of view (FOV). To tackle this issue, and as illustrated in Figure 1, each vehicle can exchange perception data with other vehicles (or road users) to extend its awareness beyond its limited individual perception.

In general, there are two ways to exchange data between the vehicles (Meuser, 2020): (a) with the help of base stations (infrastructure-based) which play the role of data forwarders between the vehicles, (b) without any help of external infrastructure (infrastructure-less) in which vehicles directly exchange data with each other. The appropriate communication type depends on the use case. Moreover, and in some use cases, a hybrid type should be applied, using both aforementioned communication types for serving a certain use case. Regardless of the communication type, each vehicle encapsulates its perception data in Collective Perception Messages (CPMs) (ETSI, 2019) and sends them to other vehicles. This perception data comes with different detail levels (Pilz et al., 2021). The receiving vehicle then fuses the received objects with its local environment model. This includes different processing tasks, like coordinate transformation (Shan et al., 2021), and enriches the vehicle's perception of the surrounding environment. This enhanced perception can be exploited in many other safety and efficiency applications in vehicular networks.

Although collective perception sounds promising, and applying it can be essential for enhancing safety and efficiency, previous works, however, show that it comes with different kinds of challenges. In the next section, we introduce some of the main collective perception challenges.

Ahmad Khalil, Tobias Meuser, Yassin Alkhalili, Antonio Fernandez Anta, Lukas Staecker, Ralf Steinmetz. *In Proceedings of the 8th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2022), pages 346-352*

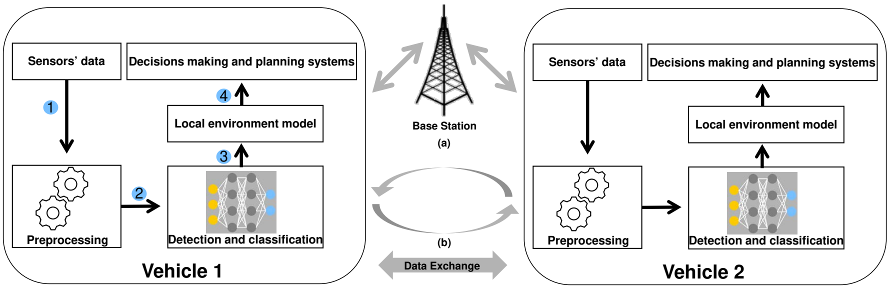


Figure 1: Illustration of the main components of collective perception. (a) Two vehicles are exchanging perception data with the help of a base station (infrastructure-based). (b) Two vehicles are directly exchanging perception data (infrastructure-less).

2.1 Collective Perception Challenges

Collective perception is a complex task and comes with many challenges. While performing collective perception, issues can occur in its all stages, *sensing*, *communicating*, and *fusing* (Pilz et al., 2021). This introduces major computational and communication difficulties. In the following, we will focus on two main challenges of collective perception: we start with illustrating the issue of having statically trained DNNs and making the case for context-specific DNNs. Then, networking issues that could occur while exchanging the heavy perception data are presented.

2.1.1 Statically trained DNNs

One pivotal aspect which affects the vehicle's scene understanding capability is how the DNNs used for object detection perform. The main weakness of the currently used DNNs comes from the fact that they are statically trained. Moreover, the data sets used for the training are commonly context-agnostic, which means that they are not tailored to specific situations (context-specific). These situations can be, e.g., location-specific (geographic), time-specific, or weather-specific. Using context-agnostic data sets for statically training the DNNs reduces their detection ability in certain situations, which may cause safety disasters (Dickson, 2020). On the other hand, to improve the performance of the DNNs, a massive amount of context-agnostic training data should be gathered, and a long training time is required. The training time and the amount of data required for the generalized DNNs scale exponentially with the required accuracy of these models (NEIL et al., 2021). Moreover, these DNNs may not reach, in the end, the intended performance in specific situations. This highlights the interest of evolving to DNNs that are dynamically trained and more specific to the different kinds of situations. In Sections 3 and 4, we will describe an approach on training and deploying situation-specific DNNs by employing the emerging Federated Learning (FL) concepts in the collective perception process.

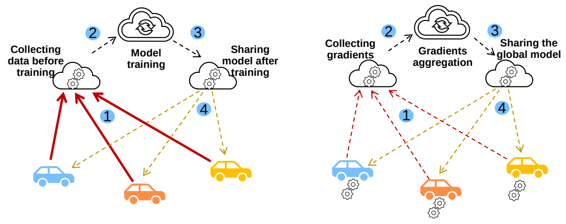
2.1.2 High-volume Data Exchange

In their work, Fukatsu et al. (Fukatsu and Sakaguchi, 2019) measured the data rate required for vehicles equipped with 3D LiDAR trying to execute collective perception for overtaking. Their study showed that even with vehicles' velocity of 50 km/h, the required data rate can easily reach around 6 Gbps if only raw data was exchanged. Considering that this data rate is generated only by a single perception application, thus, running more vehicular applications which require cooperation and data exchange between the vehicles will introduce serious challenges for the vehicular communication network. This emphasizes the importance of reducing the amount of transmitted data, either by reducing the number of exchanged messages or by reducing its volume. In Section 4, we propose Situational Collective Perception (SCP) to reduce the data exchanged, making it suitable for improving vehicles' perception.

3 FEDERATED LEARNING IN VEHICULAR NETWORKS

In this section, we provide background on the application of Federated Learning (FL) approaches in vehicular networks. We employ FL as a basis for our Situational Collective Perception (SCP) approach. FL is a promising machine learning model-training approach which has been introduced (Brendan McMahan, 2017) as a solution for privacy issues of the conventional model training approaches (Li et al., 2020). With FL, instead of collecting the training data from clients in the cloud, data is kept locally at each client. A selected set of clients train their models locally with their local data and then send the updated models (i.e.,

Ahmad Khalil, Tobias Meuser, Yassin Alkhalili, Antonio Fernandez Anta, Lukas Staecker, Ralf Steinmetz. *In Proceedings of the 8th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2022), pages 346-352*



(a) Central training (b) FL training

Figure 2: Illustration of two dynamic detection deep neural networks training approaches, (a) shows the conventional approach, in which raw data (images, point cloud data) is transmitted to the central server to be used for training the DNN. (b) vehicles only transmit their model updates to the central server to be aggregated.

the gradients) back to the server, which aggregates the model and sends it back to the clients. With its ability to work efficiently with Non-Independent and Identically Distributed (Non-IID) data of a huge number of clients, FL has attracted the interests of many industrial fields. Recently, FL has received considerable attention in vehicular networks research (Du et al., 2020; Elbir et al., 2020). More literature is getting published in recent years highlighting the privacy-preserving features of applying FL to vehicular networks (Otoum et al., 2020; Yu et al., 2020; Yuan et al., 2021; Boualouache and Engel, 2021; Zhao et al., 2021; Ding et al., 2022). In (Vyas et al., 2020), Vyas et al. proposed an architecture based on FL for developing a model that predicts the driver stress level, and then use these predictions as a basis for a driver recommendation system.

Furthermore, FL facilitates obtaining dynamically-trained onboard Deep Neural Networks (DNNs) with a reduced volume of exchanged data. So far, there are two approaches to obtain dynamically-trained DNNs, to improve its performance over time. Figure 2 illustrates those two approaches. In the conventional DNNs training approach (Figure 2a), vehicles transmit the raw data (images, point cloud data, etc.) to a central server. The central server uses the gathered data to train the global model and transmits it to the vehicles after validating it. This approach induces heavy data volume transmission, which can lead to network congestion issues. Moreover, the transmitted data can contain sensitive information which can be leaked during transmission or while processing it at the central server (Nasr et al., 2019), which leads to privacy issues. On the other hand, and as depicted in Figure 2b, the FL DNN training approach has the potential to address such networking and privacy issues. Instead of transmitting heavy volume and sensitive data to the central server, a selected number of vehicles can participate in the training process without transmitting their local data to the server. The central server collects vehicles' model updates and aggregates them using common averaging algorithms like *federated averaging* (Konečný et al., 2016). After that, the vehicles receive the global model, which replaces their local models (DNNs).

This approach enables maintaining dynamically-trained DNNs, taking into consideration the communication network limitations as well as the privacy concerns (Barbieri et al., 2021; Nguyen et al., 2021). Nevertheless, the existing literature using FL in vehicular applications focus on generating DNNs that are trained with context-agnostic data. Making these DNNs prone to fail while detecting objects in certain situations as mentioned in Section 1. A reasonable approach to tackle this issue could be to introduce situation-specific DNNs which can be used by the vehicles in specific situations.

In Section 4, we tailor the collective perception architecture in a FL way to enable dynamically trained, and situational DNNs which lead to adaptive, and efficient collective perception.

4 SITUATIONAL COLLECTIVE PERCEPTION

In this section, we provide a comprehensive explanation of our forward-looking Situational Collective Perception (SCP) architecture as a potential solution for the issues of having onboard Deep Neural Networks (DNNs) which are statically trained with context-agnostic data.

In the following, we introduce the SCP roles, then explain the models exchanging workflow between these different roles. Next, we highlight some of the potential research directions and challenges in the area of SCP.

4.1 SCP Roles

Our proposed SCP design consists of three components (see Figure 3):

- **Vehicles or other entities participating in Collective Perception (CP)**: each vehicle communicates with appropriate edge servers to retrieve the suitable situational models, after detecting a situation change. With the increased number of vehicular applications, car manufacturers are continuously improving the cars' computational power. Thus, we assume that the vehicles have the required computational power for training and validating the received situational models. Moreover, we assume for simplicity that all vehicles have

Ahmad Khalil, Tobias Meuser, Yassin Alkhalili, Antonio Fernandez Anta, Lukas Staecker, Ralf Steinmetz. *In Proceedings of the 8th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2022), pages 346-352*

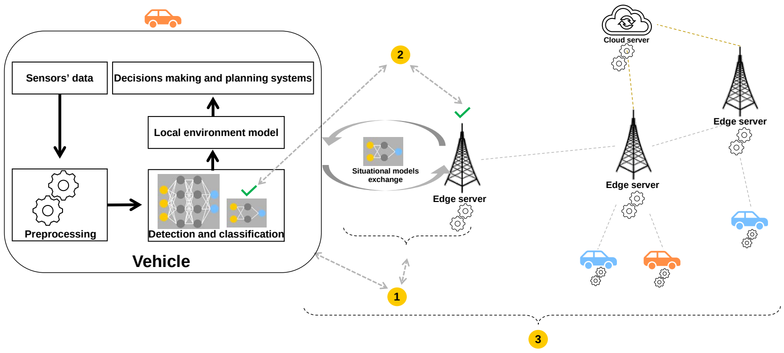


Figure 3: Illustration of three research directions in the situational collective perception.

the same sensor setup, so that they can work with copies of the same situational DNNs.

- **Edge Server**: unlike a simple base station illustrated in Figure 1 which only plays the role of data forwarder between the vehicles, we assume that the edge server has the required computation and communication power to maintain its responsibilities. The edge server is an essential part of the SCP architecture, as it is responsible for managing the different situational models, and sending the proper one to the legitimate vehicles. Besides communicating with vehicles, the edge server has to communicate with other edge servers and the cloud server in order to exchange and validate the different situational models.
- **Cloud Server**: the main task of the cloud server is to orchestrate the overall SCP processes. The cloud server is responsible for initializing the situational models and transferring them to the edge servers. Having a wide set of different situational models spread across the edge servers, the cloud server communicates with all edge servers to maximize the benefits by combining the learning parameters contained in the situational models.

4.2 SCP Workflow

To express the validity of the situational models, we introduce two different models' flags: The *training* flag indicates that the current version of the model is non-valid and vehicles can not rely on it for detection purposes. Thus, vehicles have to train the model with their local data, to eventually make it valid. On the other hand, the *valid* flag indicates that the current version of the model is valid and thus, vehicles can use it for detection purposes and train it simultaneously.
The overall process starts when the cloud server uses a pre-defined set of situations to initiate the situational models. These models then are trained initially by the cloud server with the available context-specific data sets. If no context-specific data set is found for a situational model, the cloud server skips the initial training process for that model. Later, the cloud server transmits the situational models to the edge servers, after marking the models with *training* flags, as they are not ready yet to be used for detection purposes.

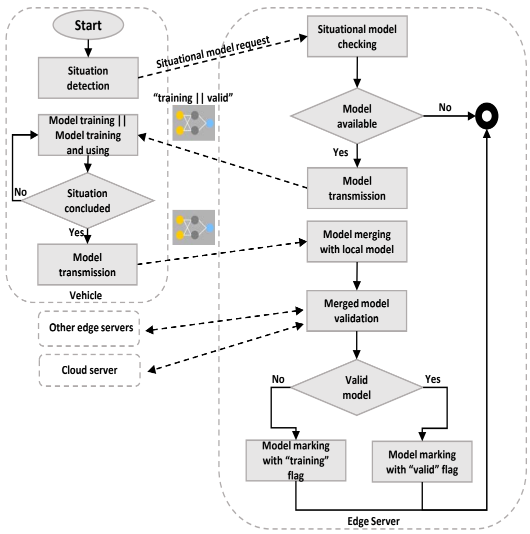


Figure 4: Illustration of situational models exchanging workflow.

After that, and as depicted in Figure 4, when the vehicle detects a new situation, it sends a situational model request to the edge server. The Edge server checks the availability of the requested situational model. If the situational model is available, the edge server transmits it to the requesting vehicle. Upon receiving the situational model, the vehicle checks the model's flag. If the model is *training* flagged, the vehicle only trains it with its locally gathered data. On the other hand, if the vehicle receives a *valid* model, it can use it for detection purposes while training it simultaneously. As soon as the vehicle detects the situation change, it re-transmits the model back to the edge server for validation. When receiving a model from the vehicle, the edge server merges it with its local situational model and validates the produced model. If the validation for produced model fails (e.g., the new version is less performing than the previous one and it does not reach a specific performance threshold), the edge server marks it with the *training* flag. On the other hand, if a valid model is obtained, the edge server marks it with the *valid* flag. Finally, the edge server replaces the old version of the model with the newly produced version.

Although this workflow sounds feasible, one question that needs to be asked, however, is how the vehicles will be able to annotate their local data. In the next section 4.3, we will provide one possible technique that enables the vehicles to annotate their local data before using it for training the situational models.

4.3 SCP Research Directions and Challenges

Figure 3 reveals three possible research areas related to SCP. Each research area is associated with a set of challenges that needs to be addressed in order to obtain benefits from employing SCP in real-world applications.

We organize SCP future research directions under three headings:

1. **Utilize Situational Collective Perception to Enable Situational Models Exchange**: since the idea of using situation-specific models is not introduced in the past in the vehicular applications area, it is important to evaluate its applicability in real-world scenarios. Moreover, during the models exchanging process (see Figure 4), the vehicles have to annotate their locally gathered data before using it for training the situational models. This is considered to be one of the main challenges in Federated Learning (FL). A possible way to handle this challenge is by using Active Learning (AL) approach (Ahmed et al., 2020). With AL, vehicles can use pre-trained models for selecting and annotating the training data with an expected level of uncertainty. This annotated data can be used then to train the situational models. Moreover, comprehensive studies are required to evaluate if the SCP approach outperforms in specific situations the conventional approach of having statically trained models using context-agnostic data. In addition, the performance of SCP relies mainly on the number of participating vehicles. Thus, enabling efficient SCP in rural areas could be a difficult task.

2. **Situational Models' Validation**: it is important to bear in mind that vehicular applications are very safety and security-sensitive. Intensive validation mechanisms need to be developed to ensure the validity of the different situational models. Validation is not only required from the edge server-side but also the vehicles must validate the received models. Vehicles may require to perform a validity check on the *valid* models received before relying on them for detection. On the other hand, edge servers are required to validate the situational models continuously, with the help of other edge servers and the cloud server. It is crucial though to develop validation criteria, defining the most important parameters which make the model either marked with *valid* or *training* flag. Another possible area of future research would be to investigate on how to ensure that SCP architecture is resistant against the anomalous nodes (vehicles, or edge servers).

3. **Situational Collective Perception at Large Scale**: perhaps one of the most powerful features of SCP comes from having a huge collection of multiple situational models spread across a wide area. However, in order to maximize the benefits, future research has to develop intensive mechanisms to deal with the heterogeneity at a large scale (different models for diverse situations). Another possible area of future research would be to investigate on how to manage the situational models' exchange between the different entities, to ensure robustness.

5 CONCLUSION

To detect their surroundings, vehicles employ statically trained onboard DNNs. These DNNs are prone to detection failures in some situations, as they are commonly trained with context-agnostic data. In this work, we introduced the forward-looking SCP to tackle these issues. Inspired by the FL approach, we tailor the collective perception architecture to enable dynamically trained and situation-specific DNNs. We aim to facilitate enabling adaptive, and efficient collective perception in future vehicular networks. We highlighted three possible SCP research areas and emphasized some of its expected challenges.

Ahmad Khalil, Tobias Meuser, Yassin Alkhalili, Antonio Fernandez Anta, Lukas Staecker, Ralf Steinmetz. *In Proceedings of the 8th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2022), pages 346-352*

ACKNOWLEDGEMENTS

This work has been funded by the German Research Foundation (DFG) within the Collaborative Research Center (CRC) 1053 MAKI.

REFERENCES

3GPP (2016). Study on enhancement of 3gpp support for 5g v2x services (release 15). *Tech. Report TR*, 22.

Ahmed, L., Ahmad, K., Said, N., Qolomany, B., Qadir, J., and Al-Fuqaha, A. (2020). Active learning based federated learning for waste and natural disaster image classification. *IEEE Access*, 8:208518–208531.

Barbieri, L., Savazzi, S., and Nicoli, M. (2021). Decentralized federated learning for road user classification in enhanced v2x networks. In *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6. IEEE.

Boban, M., Kousaridas, A., Manolakis, K., Eichinger, J., and Xu, W. (2018). Connected roads of the future: Use cases, requirements, and design considerations for vehicle-to-everything communications. *IEEE vehicular technology magazine*, 13(3):110–123.

Boualouache, A. and Engel, T. (2021). Federated learning-based scheme for detecting passive mobile attackers in 5g vehicular edge computing. *Annals of Telecommunications*, pages 1–20.

Brendan McMahan, Daniel Ramage, R. S. (2017). Federated learning: Collaborative machine learning without centralized training data.

Dickson, B. (2020). Why deep learning won't give us level 5 self-driving cars.

Ding, A. Y., Peltonen, E., Meuser, T., Aral, A., Becker, C., Dustdar, S., Hiessl, T., Kranzlmüller, D., Liyanage, M., Maghsudi, S., et al. (2022). Roadmap for edge ai: a dagstuhl perspective. *ACM SIGCOMM Computer Communication Review*, 52(1):28–33.

Du, Z., Wu, C., Yoshinaga, T., Yau, K.-L. A., Ji, Y., and Li, J. (2020). Federated learning for vehicular internet of things: Recent advances and open issues. *IEEE Open Journal of the Computer Society*, 1:45–61.

Elbir, A. M., Soner, B., and Coleri, S. (2020). Federated learning in vehicular networks. *arXiv preprint arXiv:2006.01412*.

ETSI (2019). Etsi tr 103 562.

Fukatsu, R. and Sakaguchi, K. (2019). Millimeter-wave v2v communications with cooperative perception for automated driving. In *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, pages 1–5. IEEE.

Konečnỳ, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., and Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*.

Li, T., Sahu, A. K., Talwalkar, A., and Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3):50–60.

Löffler, R. (2021). How many cars are in the world?

Meuser, T. (2020). Data management in vehicular networks-relevance-aware networking for advanced driver assistance systems.

Nasr, M., Shokri, R., and Houmansadr, A. (2019). Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In *2019 IEEE symposium on security and privacy (SP)*, pages 739–753. IEEE.

NEIL, C., THOMPSON, K., GREENEWALD, K., LEE, G., and MANSO, F. (2021). Deep learning's diminishing returns.

Nguyen, A., Do, T., Tran, M., Nguyen, B. X., Duong, C., Phan, T., Tjiputra, E., and Tran, Q. D. (2021). Deep federated learning for autonomous driving. *arXiv preprint arXiv:2110.05754*.

Otoum, S., Al Ridhawi, I., and Mouftah, H. T. (2020). Blockchain-supported federated learning for trustworthy vehicular networks. In *GLOBECOM 2020-2020 IEEE Global Communications Conference*, pages 1–6. IEEE.

Pilz, C., Ulbel, A., and Steinbauer-Wagner, G. (2021). The components of cooperative perception-a proposal for future works. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 7–14. IEEE.

Shan, M., Narula, K., Wong, Y. F., Worrall, S., Khan, M., Alexander, P., and Nebot, E. (2021). Demonstrations of cooperative perception: safety and robustness in connected and automated vehicle operations. *Sensors*, 21(1):200.

Sjoberg, K., Andres, P., Buburuzan, T., and Brakemeier, A. (2017). Cooperative intelligent transport systems in europe: Current deployment status and outlook. *IEEE Vehicular Technology Magazine*, 12(2):89–97.

Vyas, J., Das, D., and Das, S. K. (2020). Vehicular edge computing based driver recommendation system using federated learning. In *2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pages 675–683. IEEE.

Yu, Z., Hu, J., Min, G., Xu, H., and Mills, J. (2020). Proactive content caching for internet-of-vehicles based on peer-to-peer federated learning. In *2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS)*, pages 601–608. IEEE.

Yuan, X., Chen, J., Zhang, N., Fang, X., and Liu, D. (2021). A federated bidirectional connection broad learning scheme for secure data sharing in internet of vehicles. *China Communications*, 18(7):117–133.

Zhao, L., Ran, Y., Wang, H., Wang, J., and Luo, J. (2021). Towards cooperative caching for vehicular networks with multi-level federated reinforcement learning. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE.