A Distributed Learning Architecture for Semantic Communication in Autonomous Driving Networks for Task Offloading

Guhan Zheng, Qiang Ni, *Senior Member, IEEE*, Keivan Navaie, *Senior Member, IEEE*
Haris Pervaiz, *Member, IEEE* and Charilaos Zarakovitis, *Member, IEEE*

***Abstract*—Semantic communication based on machine learning (ML) techniques emerged as a new transmission paradigm that can significantly improve spectrum efficiency. It looks promising for improving the task offloading quality of service (QoS) for autonomous driving networks (ADNs) where autonomous vehicles require a significant amount of communication with the vehicle edge clouds (VECs). However, in practical ADNs, updating the ML-based semantic communication coder model is affected by various unique factors such as mobility and privacy considerations. Therefore, in ADNs, the conventional ML frameworks are not directly applicable to updating semantic communication coders. In this article, we discuss the unique challenges faced by updating the semantic communication coder in ADNs and review the existing ML frameworks. To address these challenges, we further propose a Privacy-Preserving Personalised Federated Learning (3PFL) framework for updating the semantic communication coder in ADNs. Simulation results confirm the effectiveness of 3PFL for updating the semantic communication coder in ADNs.**

I. Introduction

THE development of wireless communication systems is essential for the large-scale adoption of autonomous vehicles. As a major driving force, the vehicle edge cloud (VEC) might be considered an integral part of the 6G systems. VEC further extends cloud services to the edge of the autonomous driving networks (ADNs), e.g., roadside units, base-stations. Such a system supports the provisioning of resource-intensive applications in autonomous vehicles [1]. VEC provides computational resources to the vehicles thereby reducing service latency and energy consumption [2] while significantly improving the Quality-of-Service (QoS) of the ADNs.

The emerging automotive applications such as mobile augmented reality (AR)/ virtual reality (VR) are expected to further increase the required computational and storage resources of the autonomous vehicle. It results in autonomous vehicles increasingly offloading tasks to the VECs. Therefore, the spectrum efficiency of the ADNs should be improved to ensure the required QoS for the offloaded tasks, such as jobs of processing images or videos. Nonetheless, autonomous vehicle communication systems have been designed based on the conventional Shannon paradigm [3] and operate very close to the Shannon capacity limit. Therefore, it is an immediate need to investigate new approaches to extend spectrum efficiency beyond the conventional capacity limit to ensure the required QoS of the ADNs.

The recent development of machine learning (ML) technologies enabled the integration of semantic communication into ADNs as a promising solution for improving channel spectrum efficiency. In contrast to the Shannon paradigm that focuses on the accuracy of symbol transmission, semantic communication exploits ML to extract the actual meaning of information to reduce the transmission information quantity [4]. In semantic communication, the conventional coder is substituted by a semantic joint source-channel coder that compresses and transmits semantic information, where the coder is an ML-based Autoencoder model [5]. It thus goes beyond the Shannon capacity limit by shifting the proportion of the work to computational resources from communication and significantly increases the spectral efficiency [6].

Predictably, vehicles offloading tasks via the semantic encoder to the VEC with the semantic decoder could significantly strengthen the offloading QoS of the ADN. However, in the case of semantic communication employed in ADNs for task offloading, goal-oriented ML-based coders need to be updated in real-time for different types of task content/goals [4]. Furthermore, the training of the model requires joint participation of the encoder and decoder, i.e., the vehicles and the VEC. Therefore, real-time training of semantic coder models in this distributed semantic communication system consisting of vehicles and the VEC is a challenging task.

How to update network users' ML-based semantic coders in real-time is already considered one of the main challenges for semantic communication study. Recently, the federated learning (FL) framework for semantic communication was proposed as a potential solution [7], [8]. Nevertheless, semantic communication coders updating in ADNs introduces several unique challenges that cannot be handled by the existing FL framework. Therefore, it is of great importance to design a new real-time distributed training framework for semantic communication model updating in ADNs.

In this article, we first discuss the main challenges in semantic communication for ADNs. We then summarise the traditional ML model updating frameworks and their short-

G. Zheng, Q. Ni and K. Navaie are with the School of Computing and Communications, Lancaster University, LA1 4WA, UK (Email: g.zheng2, q.ni, k.navaie@lancaster.ac.uk).

H. Pervaiz is with the School of Computer Science and Electronic Engineering (CSEE), University of Essex, CO4 3SQ, UK (Email: haris.pervaiz@essex.ac.uk).

C. Zarakovitis is with the National Centre for Scientific Research Demokritos, Greece. (Email: c.zarakovitis@iit.demokritos.gr).

comings if applied to ADNs. Next, we propose a privacy-preserving personalised federated learning (3PFL) framework for updating the coder model in semantic communication for ADNs. The performance of the proposed 3PFL framework is also evaluated through simulations.

II. Key Design Challenges

The key challenges in updating the semantic coder model in ADNs include collaboration & coordination, mobility, model variations & personalisation, and privacy.

A. Collaboration & coordination

In a VEC's service range, multiple vehicles transmit offloading tasks via individual semantic encoders to the same VEC's semantic decoders. Different coders serve different mission-specific transmission contents. The vehicles' semantic coders need to be updated when new task content appears in the environment. Otherwise, the semantic accuracy during transmission would degrade considerably, thus reducing the QoS of offloading. However, the semantic coder is a joint source-channel coder based on ML. Updating the semantic coder requires the participation of the encoder and the decoder on both the vehicles' side and the VEC's side. Joint training of multiple vehicles and VEC together is challenging. Hence, in updating the models, effective coordination and cooperation in the training process are essential.

B. Mobility

At the time the coder models require updating, the vehicles should interact and cooperate with the VEC for training. However, due to the traffic situations and circumstances of the vehicles, they are likely to be located in different VECs' service ranges moving at various speeds. Therefore, some vehicles might depart the service range of a VEC during the training and thus not participate in the full training process. This may result in wasted computational and communication resources for these vehicles.

C. Model variations & personalisation

VECs cannot independently develop personalised semantic decoders for all autonomous vehicles in such dynamic environments. Because the vehicles would use their private semantic encoders to offload their tasks by transmitting to their corresponding VEC's semantic decoder. Updating different encoder models for various vehicles at the same time significantly increases the VEC computational workload. In addition, storing the corresponding decoder models for each vehicle imposes a storage burden. It is therefore challenging to train a personalised encoder for each vehicle and a joint VEC semantic decoder for the vehicles.

D. Privacy

For each vehicle, the training data may include sensitive information and thus its privacy should be preserved. Furthermore, for a vehicle, the privatised encoder determines the accuracy of all transmitted data. Vehicles' encoder models are also typically trained by vehicles consuming their internal resources and vehicles are not always comfortable sharing models publicly. The ML-based semantic encoder thus should also be considered one of the most important pieces of private data in the vehicle. Hence, the privacy of the semantic encoder must be preserved.

III. Conventional learning framework

We introduce two potential ML frameworks for updating the semantic coders in ADNs and discuss their constraints.

A. Central learning

Central learning (CL) is a collaborative learning approach developed based on the conventional approach of training neural networks on a single server. The training data from distributed users are collected by a central server. Subsequently, all training data on the central server is integrated and used as input to jointly train a machine learning model. The trained model is then returned to the participating users. Since in CL, the training data are trained directly by the machine learning model, it is therefore capable of obtaining higher accuracy relative to other distributed learning methods. In ADNs, the VEC with powerful computational capacity acts as a central node, collecting training data from the vehicles in the service area and executing the CL. The semantic encoder model is returned to the vehicles after the training is completed. As the training process is performed on the VEC, the vehicle computational load is therefore reduced.

However, the training data and encoder models from autonomous vehicles are private. Hence they might not always prepare to share them with a third party. Nevertheless, in CL, the vehicles' semantic encoders have the same model and model weights. Furthermore, the transmission of training data from vehicles to the VEC may introduce new overheads. For instance, if the semantic coders' service for such data is interrupted, the vehicles would have to choose the traditional transmission method, hence increasing the service delay. Such a large amount of training data might also cause network traffic congestion.

B. Federated learning

Federated learning (FL) is a distributed learning framework for collaborative training [9]. In each training epoch, distributed users first train the entire model on the user side using their individual training data and then upload the model weights to a central server for aggregation. The aggregated model is then sent back to the participating users. This enables individual clients to keep their private training data locally, hence preserving their data privacy and avoiding the problems associated with centralised data collection. Therefore, in an ADN with the FL framework, the vehicles need to be independently trained and then their models are uploaded to the VEC for aggregation.

As mentioned above, using the FL framework in ADNs is unable to address some of its unique challenges. For instance,

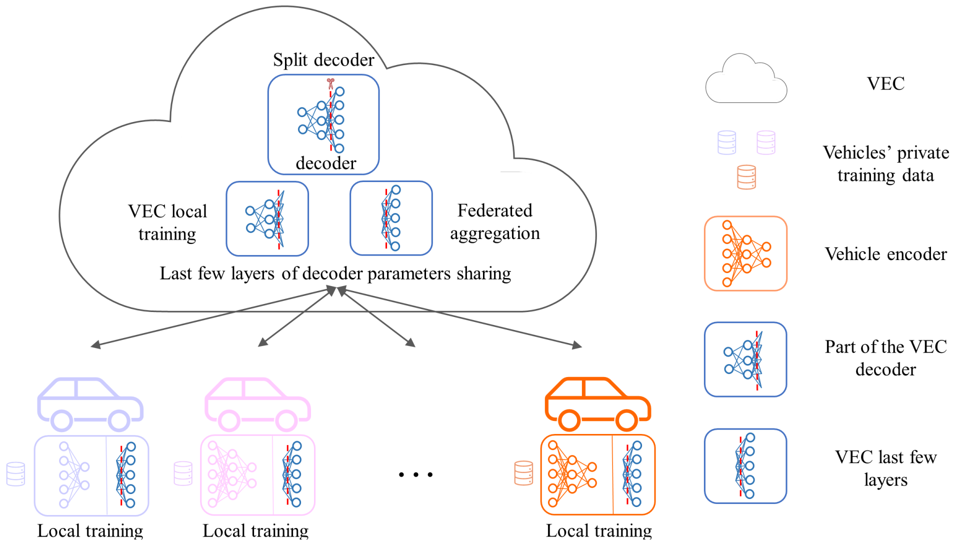

Fig. 1: The 3PFL framework.

FL cannot proceed where different vehicles have different types of encoder models. Because personalised model weights cannot be federated aggregated due to various model sizes. Furthermore, for each vehicle, FL needs to share its individual trained encoder models. Shared model parameters, however, may also result in privacy leakage as well as a great security risk as they could be utilised to infer private data [10]. Vehicles may be reluctant to share their models to prevent the encoder model exposure and the data being inferred. Therefore, the design of an effective training framework addressing the above four key design challenges for updating the semantic coders in ADNs remains an open problem.

IV. The 3PFL Framework for Semantic Communication-Based ADNs

In this section, we propose a novel framework, 3PFL, for autonomous vehicles' semantic communication that addresses the abovementioned challenges.

A. The proposed 3PFL framework components

In the proposed framework, the coder model is split into three components as follows (Fig.1):

1) Vehicles' Encoder Side: The encoder models of autonomous vehicles are locally stored for training and updating. Nevertheless, the training data and encoder models are not transmitted to the VEC. The vehicles are trained based on private data and models, enabling preserving the privacy of the vehicles and further facilitating personalisation.

2) Part of the VEC Decoder Side: Inspired by the split learning in [11], in 3PFL, semantic coders are not trained in the same location. Instead, the semantic coders are first split according to the division of encoders and decoders. The decoder is further split and the majority of the decoder's model part is kept on the VEC for training.

3) VEC Last Few Layers: The last few layers of the split decoder are transmitted by the VEC to the autonomous vehicles involved in the training. They are returned to the VEC for efficient semantic communication after completing the training.

B. Model training process

Fig. 2 illustrates an epoch of the training process in 3PFL consisting of the following five steps. The training process is performed by vehicles and the VEC together.

1) Step 1: The vehicles in each VEC service region calculate their expected residency time according to their velocity and location. After coordinating with the VEC, it is decided whether or not to engage in the updating of the semantic communication coder model with other available vehicles and the VEC. The VEC then transmits the last few layers of the semantic decoder model to the available participating vehicles.

2) Step 2: The participating training vehicles use their training data as the model inputs and train the encoder using forward propagation. The outputs obtained by the encoder are transmitted to the VEC.

3) Step 3: After receiving all the training data that has gone through step 2, the VEC takes the information received as its input. Using this input, the previous layers of the semantic decoder continue the training using forward propagation. The subsequent outputs are then sent to the respective vehicles involved in the training.

4) Step 4: After receiving the information from the VEC, vehicles train the last layers of the decoder. The finalised results are then compared with the vehicle's training data to obtain the loss value. The training network is trained once using back-propagation based on the opposite direction of the same path.

5) Step 5: At the end of vehicle back-propagation, the vehicle transmits the last few layers of the semantic decoder to the VEC for federated aggregation. The aggregated model is then returned to the individual vehicles involved in the training for the next training epoch.

In the last epoch, the VEC transmits the last few layers of the semantic decoder to the vehicles after completing the federated aggregation. The split decoders are integrated again to provide high-quality semantic communication services to the vehicles within the service range.

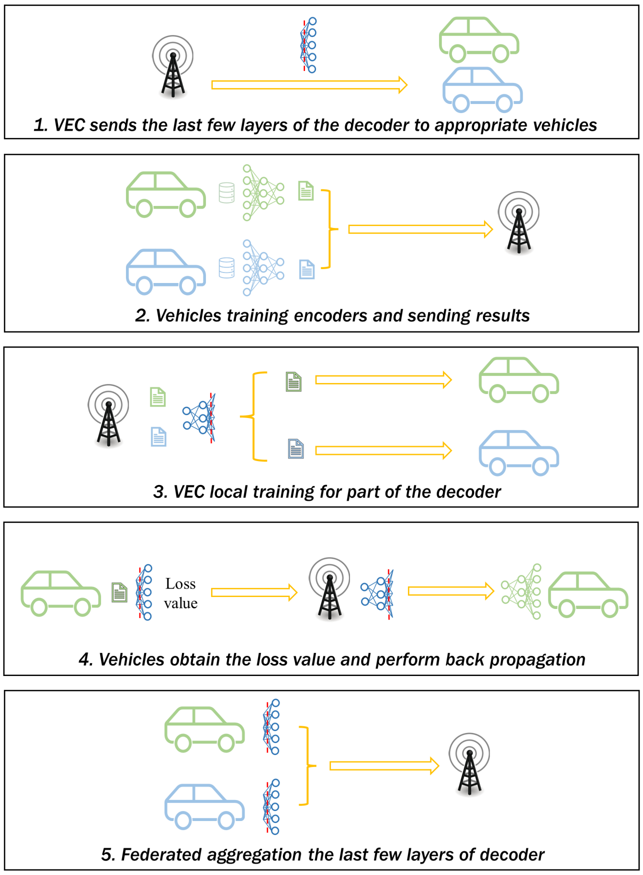

Fig. 2: The training process of 3PFL.

In terms of computational and communication overhead, we utilise the FL framework as an example for comparative evaluation. Different from the FL, in the proposed 3PFL, the personalised encoder always remains on the vehicle and the last few layers of the decoder need to be trained on the vehicle. For computation overhead, the vehicle intuitively reduces the computation of the previous layers of the semantic decoder compared to FL. For communication overhead, at each epoch, the vehicle and VEC only interacted with the last few layers of the decoder and the compressed training data from the personalised encoder. Therefore, the 3PFL has less communication overhead in case of the compressed data bits transmitted in each epoch are smaller than the data bits of the un-federated aggregation layers. Otherwise, FL has less communication overhead.

Splitting and federated aggregation, enable the training of different personalised encoder models and a jointly trained decoder model through collaboration between vehicles and VEC. Furthermore, during the entire training process, the training data and the semantic encoders are preserved on the autonomous participating vehicles. The vehicles are thus not exposing private data directly. It also reduces the risk of privacy leakage and security risks. This is due to that it enhances the privacy of the training model and thus the raw data is hard to be indirectly inferred. Based on the above discussion we argue that 3PFL can address the challenges in updating the semantic communication coder model.

V. Performance Evaluation

To evaluate the performance of the proposed 3PFL, we consider an autonomous vehicle object/image recognition offloading scenario with 10 trainable participating vehicles. Autonomous vehicles transmit images to the VEC through semantic encoders trained by various training frameworks. The VEC performs object/image recognition of the received images. The transmitted images' recognition accuracy is compared with the recognition accuracy of the pre-transmission images to derive the performance of the different semantic encoders for image offloading. As a performance metric of various semantic communication coders, we consider the semantic communication coder "accuracy" as the proportion of the received object/image recognition accuracy to the pre-transmission object/image recognition accuracy. We compare 3PFL with two baseline approaches including CL and FL. It is also assumed that all autonomous vehicles utilise the same convolutional autoencoder semantic coder model based on [12]. It is to ensure that CL and FL can be utilised successfully with the same trained model in ADNs. We use a standard image dataset, namely, CIFAR 10 [13], in our experiments. This dataset consists of 50000 images for the training set and 10000 images in the test set. The images have $3 \times 32 \times 32$ pixels and both the training and test sets include 10 different image classes for semantic extraction. The transmitted image data is recognised by the trained Densenet 201 [14] which is a commonly used ML algorithm for object/image recognition.

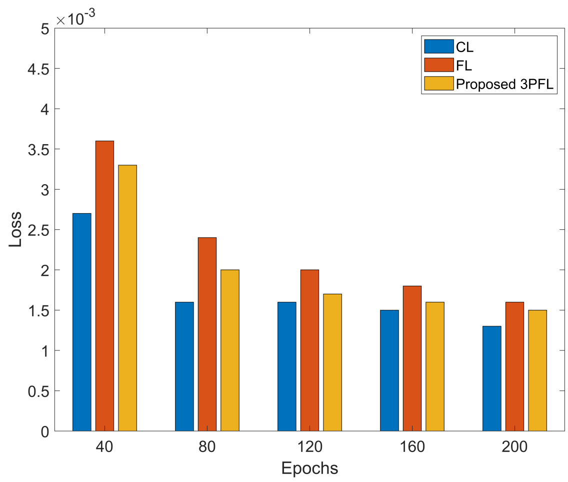

Fig. 3: Convergence speed of different ML frameworks.

Fig. 3 shows the training convergence speed of the semantic coder for various ML frameworks. By increasing the number of epochs, it is seen that all three frameworks gradually converge. Further, CL has the fastest convergence speed and the lowest loss value. It is because CL increases the accuracy by collecting all the data and training them uniformly. The proposed 3PFL framework converged almost as fast as the FL framework but with slightly lower loss values.

TABLE I: Privacy leakage of different ML frameworks

CL	FL	Proposed 3PFL
100%	45.2%	0.26%

Table 1 demonstrates the contributions of different models in preserving data and model privacy during coder updating transmission. We utilise a general data privacy leakage evaluation model from [15]. The training data, encoder model data and decoder model data are all considered. Moreover, the training data and training model are considered to have the same privacy significance. The results are normalised. It can be observed that the CL exposes all of the privacy content during transmission because it requires the sharing of all the data and the final model. The FL keeps the data local but also results in a partial privacy leakage as it exposes the entire final model parameters. Our proposed 3PFL privacy leakage value, however, is much smaller than other frameworks. This is because our proposed model maintains the training data, the users' personalised encoder models, and part of the decoder model locally which are not shared/exposed. Further, It also guarantees the privacy of the users' training data and the encoder model by preventing them to be indirectly inferred. The proposed 3PFL thus enhances the privacy of the users' training data and the encoder model's privacy and personalisation.

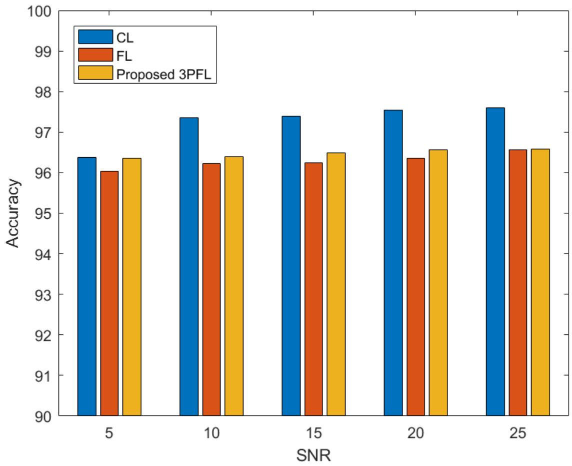

Fig. 4: Accuracy of different ML frameworks in case of AWGN.

Fig. 4 and Fig. 5 illustrate the impact of testing accuracy on various frameworks in different communication environments. This embodies the key performance indicator (KPI) of the effectiveness of the semantic communication encoder in executing semantic communication [7]. We set the semantic encoder output layer to 10 neurons. In Fig. 4, we evaluate the influence of additive white Gaussian noise (AWGN) on the trained semantic communication coders. It is seen that as the signal-to-noise ratio (SNR) increases, the performance of the coders slightly increases. The CL achieves the highest performance for different SNR settings. Furthermore, our proposed 3PFL performance is slightly higher than FL. In Fig. 5, the effect of Rayleigh fading is shown. It can be observed that the semantic coder model trained under different frameworks increases with increasing SNR same as Fig. 4. Moreover, the accuracy of our proposed framework continues to perform excellently after CL in various SNR environments. Nevertheless, it is notable that in contrast to 3PFL applicability to a variety of vehicle offloading environments, the baseline frameworks are only implementable in extremely special scenarios. For instance, in the case of vehicles having the same encoder models and without privacy considerations. The above results confirm that the proposed framework is effective for semantic communication applications in autonomous driving and either outperforms (or is equally efficient as) the baseline frameworks while preserving the privacy of the vehicles' data.

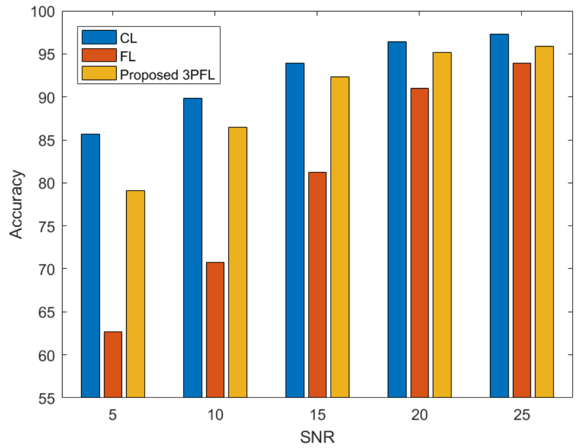

Fig. 5: Accuracy of different ML frameworks in case of Rayleigh fading.

VI. CONSLUSION

We investigated the application of semantic communication in the VEC for offloading tasks by autonomous vehicles. We discussed the technical challenges in updating the semantic communication model in ADNs. We then briefly reviewed the existing solutions to these challenges and discussed two promising ML frameworks. It was found that the existing ML frameworks are unable to satisfy the requirements of the ADNs. We then proposed a novel 3PFL framework for updating the semantic coders in ADNs. In 3PFL a coder is split into three parts for training and only requires the last few layers of the decoder for federated aggregation. Simulation results confirmed the effectiveness of 3PFL for updating the semantic communication coders in ADNs.

REFERENCES

[1] S. S. Shinde, A. Bozorgchenani, D. Tarchi and Q. Ni, "On the Design of Federated Learning in Latency and Energy Constrained Computation Offloading Operations in Vehicular Edge Computing Systems," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 2, pp. 2041-2057, Feb. 2022.

[2] J. Wang, J. Hu, G. Min, Q. Ni and T. El-Ghazawi, "Online Service Migration in Mobile Edge with Incomplete System Information: A Deep Recurrent Actor-Critic Learning Approach," *IEEE Transactions on Mobile Computing*, 2022.

[3] C. E. Shannon and W. Weaver, The Mathematical Theory of Communication. The University of Illinois Press, 1949.

[4] Z. Qin, X. Tao, J. Lu, and G. Y. Li, "Semantic communications: Principles and challenges," Jun. 2022, [online] Available: http:// arXiv:2201.01389.

[5] Q. Lan et al., "What is Semantic Communication? A View on Conveying Meaning in the Era of Machine Intelligence," *Journal of Communications and Information Networks*, vol. 6, no. 4, pp. 336-371, Dec. 2021.

[6] C. Chaccour, W. Saad, M. Debbah, Z. Han, and H. V. Poor, "Less data, more knowledge: Building next generation semantic communication networks," Nov. 2022. [Online]. Available: https://arxiv.org/abs/2211.14343.

[7] G. Shi, Y. Xiao, Y. Li and X. Xie, "From Semantic Communication to Semantic-Aware Networking: Model, Architecture, and Open Problems," *IEEE Communications Magazine*, vol. 59, no. 8, pp. 44-50, August 2021.

[8] Z. Qin, G. Y. Li and H. Ye, "Federated Learning and Wireless Communications," *IEEE Wireless Communications*, vol. 28, no. 5, pp. 134-140, October 2021.

[9] H. B. McMahan, E. Moore, D. Ramage, and B. A. y Arcas, "Federated learning of deep networks using model averaging," 2016. [Online]. Available: http://arXiv:1602.05629

[10] S. Truex et al., "A hybrid approach to privacy-preserving federated learning" in arXiv:1812.03224, 2018, [online] Available: http://arxiv.org/abs/1812.03224.

[11] P. Vepakomma, O. Gupta, T. Swedish, and R. Raskar, "Split learning for health: Distributed deep learning without sharing raw patient data," 2018, [online] Available: http://arXiv:1812.00564

[12] E. Bourtsoulatze, D. Burth Kurka and D. Gündüz, "Deep Joint Source-Channel Coding for Wireless Image Transmission," IEEE Transactions on Cognitive Communications and Networking, vol. 5, no. 3, pp. 567-579, Sept. 2019.

[13] A. Krizhevsky, "Learning multiple layers of features from tiny images," Univ. Toronto, Toronto, ON, Canada, Tech. Rep. TR-2009, 2009

[14] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q Weinberger, "Densely connected convolutional networks," *CVPR*, pages 4700–4708, 2017.

[15] R. Xing, Z. Su, N. Zhang, Y. Peng, H. Pu and J. Luo, "Trust-Evaluation-Based Intrusion Detection and Reinforcement Learning in Autonomous Driving," *IEEE Network*, vol. 33, no. 5, pp. 54-60, Sept.-Oct. 2019.

Guhan Zheng is a Ph.D. candidate in communication systems at the School of Computing and Communications, Lancaster University, Lancaster, U.K. His research interests include machine learning, semantic communication, dynamic communication networks and 3D communication systems.

Qiang Ni is a professor at the School of Computing and Communications, Lancaster University, Lancaster, U.K. His research areas include future generation communications and networking, including green communications/networking, millimeter-wave wireless, cognitive radio systems, 5G/6G, SDN, cloud networks, edge computing, dispersed computing, Internet of Things, cyber physical systems, artificial intelligence/machine learning, and vehicular networks. He has authored or coauthored 300+ papers in these areas.

Keivan Navaie is a professor at the School of Computing and Communications, Lancaster University, Lancaster, U.K. His research interests include distributed cloud and edge computing, connectivity resilience in cyber-physical systems and IoT, applications of machine learning and artificial intelligence, and cognitive communications networks.

Haris Pervaiz is an associate professor at the School of Computer Science and Electronic Engineering, University of Essex, UK. His research interests include 5G & beyond, green communication, semantic communication, AI/ML based networking, mmWave communication, energy & backhaul energy efficiency and green HetNets. He is an associate editor for IEEE Access, Wiley ETT, Wiley ITL, IET Networks and IET Communications. He also organised various workshops and special sessions as lead workshop chair in various IEEE flagship conferences.

Charilaos Zarakovitis is a senior researcher at the Institute of Informatics and Telecommunications (MNL), National Centre for Scientific Research "Demokritos", Greece. His research interests include quantum neural networks, machine and deep learning, green communications modelling, bioinspired and game-theoretic decision-making, cognitive radios, network virtualisation, statistical signal processing, and convex optimisation analysis.