Application of Edge Intelligence and Vehicle-To-Vehicle Communication in Next-Generation Autonomous Driving

Zhengkun Sun[1, †, *], Lechen Zha[2, †], Yanhao Zhou[3, †]

[1] Hongwen School Qingdao, Qingdao, 266000, The People's Republic of China

[2] Suzhou Foreign Language School, Suzhou, 215000, The People's Republic of China

[3] Shenzhen Foreign Languages School, Shenzhen, 518000, The People's Republic of China

[†]These authors contribute equally.

* Corresponding Author Email: steven.sun@hongwenfeh-qd.com

Abstract. Self-driving vehicles have been one of the most promising technologies in the 21st century; however, today's self-driving is way behind optimal autonomous driving where the car can take complete control. This is mainly due to a lack of accuracy and speed of critical car operation decisions made by the electronic control unit from artificial intelligence learning. Using data learning, the accuracy of the calculations is strongly related to the amount of environmental data input; nevertheless, this also proposes that current onboard processors are not powerful enough. This article discusses future solutions such as vehicle-to-vehicle communication and edge intelligence, elaborating on their mechanisms involving artificial intelligence, edge computing, and the internet of vehicles. Moreover, system and mathematic models related to vehicle-to-vehicle communication and edge intelligence are summarized; together with challenges and solutions further in the article. Overall, this article has attempted to provide a theoretical basis for a more profound exploration of deeper autonomous driving techniques and possible ways to achieve so.

Keywords: Edge Intelligence, Vehicle-to-vehicle Communication, Autonomous Driving.

1. Introduction

It often prospects that in the near future, autonomous driving techniques will fully replace human control. However, as it may seem achievable with current technologies, it is undeniable that the current data learning (DL) algorithms and models requires an immense data processing ability that is unachievable in the on-board electronic control unit (ECU) or the cooperative adaptive cruise control (CACC) system. Therefore, to achieve this, today's self-driving vehicles are introduced to cloud computing, Google Maps for instance. Despite the effective and accurate decisions made by such super-computers, the bandwidth of data transmitted must be exceptionally high together with a very significant latency; this is unacceptable for autonomous driving when internet speeds are limited and surrounding data alters significantly over a short time. Therefore, technologies such as the internet of everything (IoE), vehicle-to-vehicle (V2V) communication, short-range communication protocols such as IEEE 802.11p, faster cellular network protocols such as beyond 5th generation (B5G) or 6th generation (6G), edge computing, and artificial intelligence (AI) are all newly proposed and convincing technologies. In this article, the combination of V2V, 6G, and edge intelligence in the application of autonomous driving vehicles (AVs) will be discussed.

Currently, the Society of Automotive Engineers (SAE) defines 6 levels of driving automation ranging from levels 0-5 [1], these levels can be the standard for defining the AV's ability to self-drive. Furthermore, these levels had adapted by the U.S. Department of Transportation where level 0 stands for full manual and level 5 for fully autonomous, as shown in figure 1.

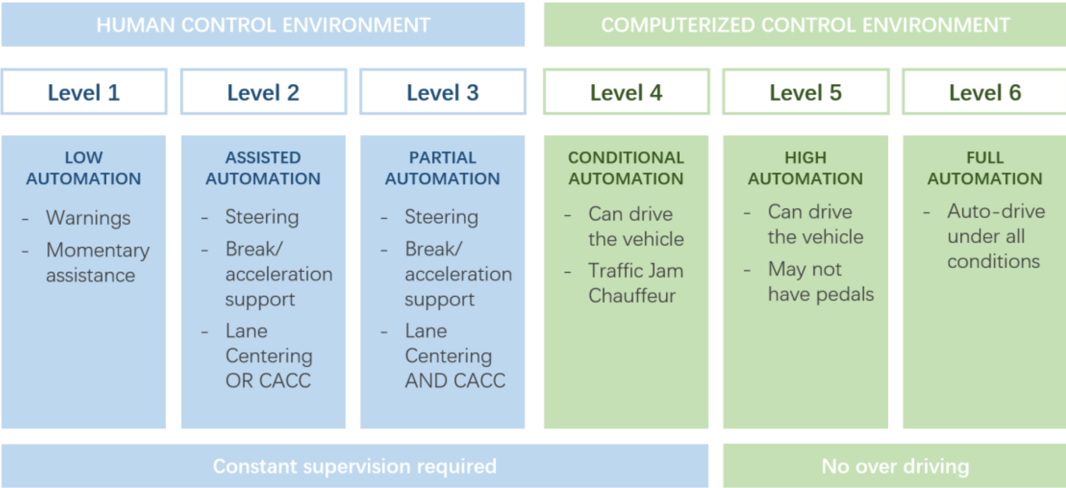

Figure 1. An illustrative level of autonomous driving ability ordered from levels 0-5 [2]

Therefore, this paper will focus mainly on the theoretical assumption, modelling, and comparisons between the current proposed technologies in attending to reach high automation levels such as 4 or even 5. Specifically, this article has summarised the principles of edge intelligence (EI) and internet of vehicles (IoV), the applications of EI, and mathematical system models including task models, communication models, and calculation models of V2V. Finally, this article has provided challenges faced recently, future aspects, together with possible solutions proposed by scholars. Ultimately, the main goal of this paper is to provide a theoretical basis for further deepening exploration of autonomous driving techniques in the future.

2. Overview of edge intelligence technology

2.1.The basic principle of edge intelligence

Edge computing technology and AI are combined to form the term "edge intelligence", which aims to provide technical improvement in artificial services. It aims to analyse and aggregate data information close to the network capture points. With edge computing, different types of data, which are originally processed at the central point of the system, can be handled by remote and decentralized system nodes. While the main ability of AI is to realize the intelligence of computers and enable the computers to work and think just like human brains. After combing these two technologies, in general, EI is considered to be a paradigm by running AI algorithms locally on an end device, with the information acquired on that device, by some organizations and presses [3]. To explore the working patterns and detect the unusual data varying trend that may happen in various situations, such as the flow of traffic and temperature trends in a specific region, the sensed data will be learned quickly and automatically through AI techniques.

2.2.The key technique of edge intelligence

2.2.1. Edge computing technique

Edge computing supports EI mainly by providing more complete data calculations, and storage of it, and pushing the information analysis close to the users. From Figure 2, it can be seen that data is delivered from edge devices like monitors, mobile devices, and televisions, and then gathered to edge nodes. Edge nodes are also called "gateway nodes" or "edge communication nodes". They are computers that act as end-user communication portals in cluster computing, which includes master nodes and workers (allocating tasks and executing tasks respectively) [4].

By setting an edge server between the end devices and the central online calculation centre, the potential in network edge points can be improved, which can be represented by the decrease in the data reaction back delay. The setup of edge servers will release the pressure in the central handle system, as it can store the data in a short term and equips with the same analysis ability as the central system. While there still exist some flaws in dealing with massive data information, due to the relatively imperfect computing ability compared with the central system. The data recorded during a long period with a huge amount is better for processing in the central processing server. In some specific situations like automatic driving, which require the digit results immediately to carry out the next step, by using the central cloud analysing system, it will be late when sending the data back due to the long travelling distance. The flow can be seen in figure 2.

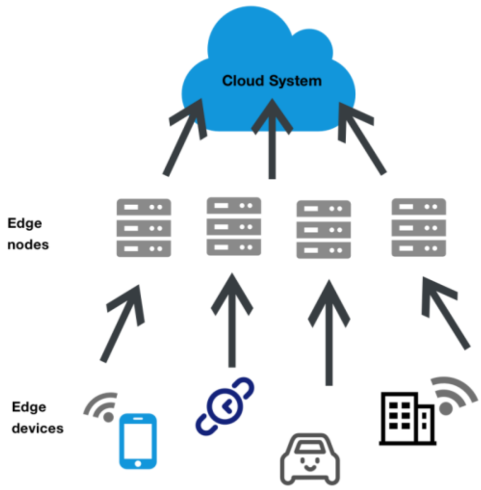

Figure 2. Illustrative model of how edge computing works

2.2.2. Artificial intelligence

AI technology is one branch of computer science, mainly researching and imitating human thoughts and behaviours. This wide field includes several aspects, such as language and image recognition and natural language processing. Applying this technique, enabling machines to perform complex tasks that normally require human intelligence. Equipping with the ability to act as human brains, a lot of issues that originally requires humans' processing strategies can be enacted through computing. As a result of AI development, recording analysis will become precise and human resources can be saved for other aspects.

2.2.3. Internet of vehicles

IoV is a media that acquire information and data from the sender units in the vehicles to conduct data analysis, which subsequently comes to the necessary conclusion to help decisions making. According to research statistics, 90% of traffic accidents are caused by drivers' errors or misjudgements [3]. If automatic driving technology and vehicular communication cooperation with the Internet could be adopted, traffic accidents caused by driving errors would be significantly reduced and traffic congestion would be greatly relieved Since 2012, with the rapid development of big data technology and the internet of things (IoT)/IoE, the first generation of IoV has made the key effect to realize autonomous driving scenarios in the future. Intuitively, IoV can be regarded as a powerful wireless sensor network (WSN) launching by itself without human intervention [5]. However, to apply IoV strict restrictions are set up which leads to some problems remaining compared with traditional WSN. Cognition and automaticity are key and particular features of both IoT and IoV.

3. The Application of Edge Intelligence in Autonomous Driving

Recently, the integrational technology of AI on edge computing has been greatly researched. The product, EI has therefore sparked a great increase in computational efficiency and accuracy in terms of autonomous driving. The use of EI in self-driving has innovated and improved 3 major applications [6], namely: object detection, traffic flow prediction, and performing intelligent decisions.

3.1.Object Detection

Today, the few cars that support self-driving work mainly on processing data simultaneously on board; however, these methods bottleneck higher levels of driving automation due to limited computing power. The collection of data on a self-driving car is based primarily on video cameras, radar sensors, and lidar sensors. Robert Bielby, in an interview, states that the rate of data created by these sensors can 'range from 3 Gbit/s (~1.4TB/h) to 40 Gbit/s (~18 TB/h)' [7]. This is a massive amount of data and is too significant to upload to the cloud, yet the computational power on cars is not effective enough to provide precise, complete, and reliable simulations [8]. Therefore, edge servers are the current best solution to achieve low latency and high-accuracy calculations.

To reduce accidents and decrease congestion, AVs are required to identify surrounding objects timely and accurate [9]. Nevertheless, executing delay-sensitive tasks with DL generally requires high computation resources and high bandwidths [6]. To improve inference accuracy and successfully address higher levels of auto-driving, transporting portions of data via B5G or 6G cellular networks with EI and V2V communication becomes a mandate.

3.2.Traffic Flow Prediction

Together with the need for fast response and high accuracy for AI, the acquisition of traffic flow being timely and accurate will also play an important role in applications such as navigation in autonomous driving. The immense environmental data is again beyond the capabilities of car storage and the system would be way too complex to satisfy the needs under current capabilities [10]. Here, 'EI creates and represents a trend of "driverless revolution on big data"' [10]. To explain, the combination of AI with advanced edge computing allows the resources to transfer to edge servers to achieve state-of-the-art performance for hierarchical learning. This enables the analysis of high-dimensional datasets (e.g., databases) to satisfy real-time requirements.

3.3.Intelligent Decision

With successful object detection and traffic flow predictions, intelligent decisions can be performed to satisfy applications such as online pathfinding to improve road safety. Critical car operations, breaking, turning, and overtaking for instance can be operated by AVs with intelligent decisions. In current practice, vehicles that support self-driving have their environmental data processed on the ECU or CACC [6], however, when the decision-making problem is taken into account with vehicle dynamics, the decisions can hardly be achieved under low delays [10]. Regarding this issue, EI can be a revolutionary breakthrough due to fine decisions created in near-real-time by a well-trained DL model [11]. Unlike cloud computing such as Google Maps, EI can significantly decrease wireless communication delay by uploading the collected data, such as vehicle conditions and surroundings data, directly to the edge server [6]. Thus, creating intelligent, timely, and reliable decisions for AVs.

4. System Modelling and Mathematical Analysis of Edge Intelligence in Autonomous Driving

To achieve autonomous driving, a system of 2 major frameworks has to be successfully utilized: the V2V network between vehicles, and Edge Intelligence.

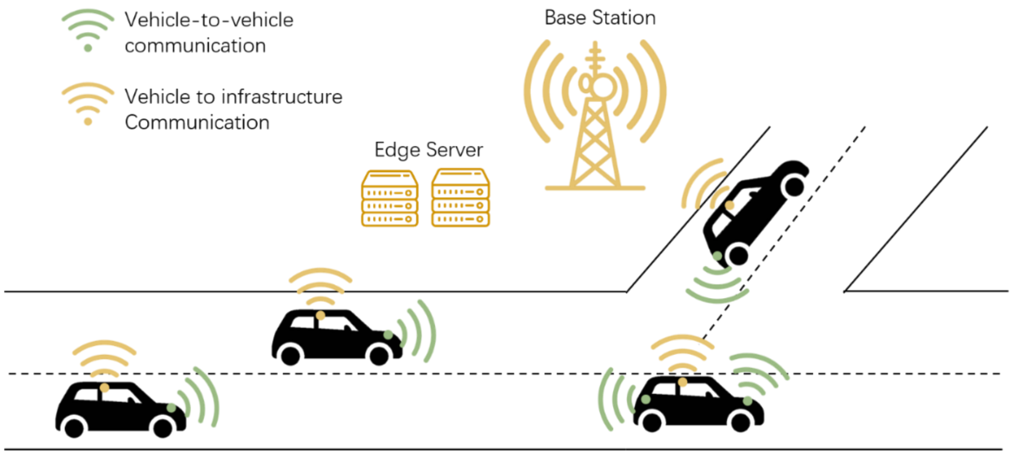

Figure 3. Illustrative model of autonomous driving network with V2V and EI in a possible scenario of an autonomous driving system.

From Figure 3, it can be concluded that autonomous driving networks can be represented by a directed graph $G = (V, E)$ [12]. Here, V is the set of vertices, representing N_v number of vehicles, and E is the set of edges in the network; representing the communication links between the vehicles in the network. However, it is worth noting that vehicles generally operate at high speeds, therefore, we consider a short period where the topology between vehicles and their communication rate remains unchanged. Lv Xinchen and Zhang Chengyu, in their article [12] have illustrated and performed such models related to multi-source data.

4.1.Task and Communication Models for V2V

To successfully classify and predict autonomous driving data via DL, the output data from a vehicle is much more massive than the data received, this ratio can range up to 10^5 [12], therefore, the bandwidth of data received can be negligible. Thus, we can define the following terms to model such data: (a) A_i as the amount of data precepted by vehicle i, (b) ρ as the complexity of the calculation, indicating the relationship between the amount of data to be processed and the required ECU clock cycles, (c) $A\rho$ as the clock cycle of the ECU to process A amounts of data, and (d) N_v as a set of vehicles in the V2V network containing $\{v_1, v_2, v_3, \ldots, v_n\}$

Furthermore, to improve the accuracy of the predictions, the distribution and reuse of environmental data created by multiple vehicles are essential. The sharing of environmental data can enable vehicles to obtain data from other nodes in the network, thereby verifying and repairing data, and ultimately improving the quality of the neural network data input and meeting the data requirements of the machine learning model. Thus, it is obvious that the operation of maximum amounts of data with limited resources has to be applied to engage V2V communication.

Based on this, 2 major communication standards have been proposed and compared, the IEEE 802.11p and cellular vehicular-to-everything (C-V2X) to be specific. The IEEE 802.11p communication standard can be applied to perform V2V communication with orthogonal frequency division multiplexing technology and has been approved by the Department of Transportation and the European Commission; to be ready for the usage of short-range communications. However, it is worth noting that it is not capable of massive data transportation which allows B5G/6G technology to take over [6].

Here, the channel transmission rate C_{ij} of the communication link between vehicle i and vehicle j is affected by inter-vehicle distance, background noise power, etc due to the Shannon-Hartley theorem shown in Equation (1). In the Shannon-Hartley theorem, C is the maximum channel capacity, B is the bandwidth of the channel, S is the average received signal power over the bandwidth, N is the average power of the noise and interference over the bandwidth, and S/N is the signal-to-noise ratio (SNR).

$$C = B\log_2\left(1 + \frac{S}{N}\right) \tag{1}$$

Therefore, if $C_{ij} = 0$, it indicates no communication between vehicles i and j because they are not in range. When they are close, however, it illustrates that communication can be established in packs of data. Nevertheless, due to a limited C_{ij}, the rate of data transfer has to satisfy equation (2)

$$d_{ij} \leq C_{ij}, \tag{2}$$

Where: d_{ij} denotes the size of data transferred within vehicles i and j. At the same time, when considering the complete and efficient use of data, it is a waste of system resources to transmit duplicate data to the same vehicle for processing. Therefore, the amount of transmitted data should not exceed the amount of data perceived by the onboard sensors, meeting the requirements of Equation (3).

$$d_{ij} \leq A_i \tag{3}$$

4.2.Calculation Models

To model EI on AV, we can continue defining the following terms: (e) F_i as the maximum computing power of the vehicle i, and (f) f_i as the current amount of data processed by the vehicle i.

The modelling consists of two conditions, namely: (1) when all data collected is analysed by the ECU, and (2) when a partial copy of the environmental data is distributed to other vehicles close by. To maximize energy efficiency, all vehicles should process local data primarily before calculating received data, this is to prevent energy loss from useless transmissions. Therefore, the data processed by vehicle i should meet Equation (4).

$$f_i = A_i + \sum_{j=1, j\neq i}^{N_v} d_{ji} \tag{4}$$

Moreover, due to limited calculation efficiency, the data calculated by vehicle i should satisfy

$$f_i \leq {}^{F_i}\!/_{\rho}\,, \tag{5}$$

thus, by substituting Equation (4) into (5), we can get

$$A_i + \sum_{j=1, j\neq i}^{N_v} d_{ji} \leq {}^{F_i}\!/_{\rho} \tag{6}$$

Each calculation requires energy consumption, if e_i stands for the energy required by vehicle i to process units of data, it is directly related to F_i and can therefore be modelled as so.

5. Challenges & Future Application of EI

5.1.Challenges

5.1.1. Trade-off between Accuracy & Latency

Edge Intelligence does predictions based on the information it gains. For a model, both its accuracy and the delay in creating it are important in determining its effectiveness. False models will lead to accidents and long delays will degrade the user's experience. However, there's a trade-off between these two vital aspects. An edge server needs time to gain information for an accurate prediction, but acquiring data means time lags. Sensors need more time to gather and transmit more information to the server. Thus, if more information is included in the analysis, then the more accurate the model is at the cost of time, and vice versa. A server has to choose between an accurate prediction that requires a lot of time or a fast prediction based on insufficient data. Photographs are analogous to this process. A picture of a higher definition requires larger storage space, while ones with poorer quality can be transmitted quickly.

One solution to this problem was proposed by Hang and his colleagues in one of their essays [13]. where they have designed a method called weight pruning to compress information while keeping the accuracy of the final inferences by servers. Traditional compressions, though alleviating the servers and reducing the latency, provide poor predictions. Weight pruning, however, keeps the accuracy of the final results. It acts on deep neural networks and is composed of three steps. The computer (i) learns which connections in the networks are vital and useful, (ii) eliminates the useless proportions of the networks, and (iii) learns the aggregate weight of connections in the final network. By keeping the vital sections of the network and reducing the others, the server can reduce the time required to process the data while staying accurate.

5.2.2. Privacy Issues

Edge Intelligence contains Artificial Intelligence, and AI needs to learn from data to work sufficiently. Here, real-world data is preferable to simulated data. Traditional servers gain information from sensors directly, sensors access data from their surroundings and upload the data to edge servers so the servers can learn them. For instance, if we frequently text a word on our phone, then our phone learns from our keystrokes. The phone learns that we are more likely to type in the word so that a specific word appears more often in the next word prediction. With the development of society, sensors will upload more and more data and some of them are sensitive. To begin with, these data can overload the servers and create privacy issues [14]. Moreover, hackers can attack both the servers and the sensors to illegally gain private information.

One solution to this problem is Federal Learning [15]. Federal Learning requires edge devices to do basic processing and computation of data. The devices make their own models based on the information they gained before uploading them to the servers. Servers then process and learn from the processed models, combine all models from the devices using V2V and reach a consensus. The servers have no direct access to sensitive information and have no risk of leaking it. Thus, Federal Learning ensures data security and reduces the workload of edge servers. It makes information anonymous and hides sensitive information. It also makes hacking behaviour less likely to succeed since it reduces the surface of the attack. Hackers can no longer gain sensitive information by attacking the servers.

5.2.Future Applications

Due to the rise of big data, the Internet now contains an increasing amount of information. The computation and transmission delay thus increases by a large amount, which is prohibitive. Services like autonomous driving require short latency, high security, and high accuracy. Else they are unreliable.

Traditional Cloud Computing can’t fulfil those requirements. Cloud servers are far from the users and all information crowds in the servers, so they have a higher delay and a higher risk of privacy leakage. Edge servers are different. Their real-time response and high bandwidth make them attractive. Computers no longer need to send information to distant cloud servers. Besides, they save energy as well since they can deliver tasks to edge servers while meeting the latency limit. Thus, edge computing can support some of the services mentioned above while improving users’ experience of currently available services. It will aid IoT and B5G/6G networks, benefiting human lives.

6. Conclusion

In this article, we have discussed the possible applications and benefits of EI, especially in automatic driving. We have explained the two basic compositions of EI together with IoV and the ways they aid automatic driving with statistics and data. Edge Computing contains less transmission of data, and Artificial Intelligence saves users' time and energy. We also explored three major applications of EI: Object detection, Traffic flow prediction, and Intelligent decision. We have then illustrated how these applications can help automatic driving with their high accuracy and low latency. Moreover, we used a model and a formula to illustrate data rates with EI and V2V when applied to automatic driving. Then, we investigated several challenges and threats posed by the new technique and possible solutions. Eventually, we discussed some future aspects and applications of the technique. How Edge Intelligence may help build IoE and support new techniques is proposed. However, we still need to do further research to overcome the challenges and optimize such technology.

References

[1] SAE Levels of Driving Automation™ Refined for Clarity and International Audience. (n.d.). Retrieved September 18, 2022, from https://www.sae.org/blog/sae-j3016-update

[2] What is an Autonomous Car? – What Self-Driving Cars Work | Synopsys. (n.d.). Retrieved September 7, 2022, from https://www.synopsys.com/automotive/what-is-autonomous-car.html

[3] Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2019). Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing. Proceedings of the IEEE, 107(8), 1738–1762.

[4] Contributor, T. (2020, September 16). edge node. WhatIs.Com. https://www.techtarget.com/whatis/definition/edge-node

[5] Min Chen, Yuanwen Tian, Giancarlo Fortino, Jing Zhang, Iztok Humar, Cognitive Internet of Vehicles, Computer Communications, Pages 58-70

[6] Yang, B., Cao, X., Xiong, K., Yuen, C., Guan, Y. L., Leng, S., Qian, L., & Han, Z. (2021a). Edge Intelligence for Autonomous Driving in 6G Wireless System: Design Challenges and Solutions. IEEE Wireless Communications, 28(2), 40–47.

[7] Mellor, C. (2020b, February 6). Autonomous vehicle data storage: We grill self-driving car experts about sensors, clouds . . . and robo taxis. Blocks and Files. Retrieved September 18, 2022, from https://blocksandfiles.com/2020/02/03/autonomous-vehicle-data-storage-is-a-game-of-guesses/

[8] Lyu X., & Yu Z. (2021). Edge Intelligence Multi-Source Data Processing for Autonomous Driving in Internet of Vehicles. Journal of Beijing University of Posts and Telecommunications, 44(2).

[9] Wang, J., Liu, J., & Kato, N. (2019). Networking and Communications in Autonomous Driving: A Survey. IEEE Communications Surveys &Amp; Tutorials, 21(2), 1243–1274. https://doi.org/10.1109/comst.2018.2888904W

[10] Xu et al., "Data-Cognition-Empowered Intelligent Wireless Networks: Data, Utilities, Cognition Brain, and Architecture," IEEE Wireless Commun.}, vol. 25, no. 1, Feb. 2018, pp. 56–63.

[11] B. Yang et al., "Computation Offloading in Multi-Access Edge Computing: A Multi-Task Learning Approach," IEEE Trans. Mob. Comp., Apr. 2020, pp. 1–1.

[12] Lyu X., & Yu Z. (2021). Edge Intelligence Multi-Source Data Processing for Autonomous Driving in Internet of Vehicles. Journal of Beijing University of Posts and Telecommunications, 44(2). K. Xiong et al., "Intelligent Task Offloading for Heteroge-neous V2X Communications," IEEE Trans. Intelligent Transportation Syst., Aug. 2020, pp. 1–13.

[13] S. Han, J. Pool, J. Tran, and W. Dally, "Learning both weights and connections for efficient neural network," in Proc. Adv. Neural Inf. Process. Syst., 2015, pp. 1135–1143.

[14] Zhu, G., Liu, D., Du, Y., You, C., Zhang, J., & Huang, K. (2020). Toward an Intelligent Edge: Wireless Communication Meets Machine Learning. *IEEE Communications Magazine*, *58*(1), 19–25. https://doi.org/10.1109/mcom.001.1900103