Deep Reinforcement Learning Based Intelligent Approach for Effective Communications in Vehicle Platoon Systems

***Abstract*—Vehicle platoons are the applications that provide greener, safer, faster, and economical transport as part of modern intelligent transportation systems. The advent of deep learning and reinforcement learning made it possible to address standard communication problems effectively using data-driven techniques. In this short research paper, we discuss the realm of vehicular platoon where deep and reinforcement learning can be applied and requires attention. We also propose an application of deep reinforcement learning to solve the resource allocation problem intelligently, addressing the specific needs of vehicle platoons. This application should pave the way for using deep reinforcement learning implements in vehicle platoon communications per se. We also discuss the scope for improvements required to well adopt deep reinforcement learning in this regime.**

***Index Terms*—Deep Learning, Reinforcement Learning, Vehicular Communication, Vehicle Platoon, Vehicular Networks.**

I. Introduction

In the last decade, on average, eighty million commercial and non-commercial vehicles were manufactured every year worldwide [1]. It indicates that a large number of vehicles have been registered for on-road operations every year. Besides, there is a significant increase in the migration of people toward urban areas. It brings about transportation-related problems such as traffic conditions, pollution, and safety for modern men in day-to-day life. Thus, it is time to update the current transportation systems accordingly.

Nonetheless, vehicles are becoming sophisticated and more conscious of the surrounding environment with time. It leads to the requirement for the next stage of connectivity, paving the way for connected vehicle technologies. It includes technologies and standards like IEEE 802.11p, which is the base for dedicated short-range communication in the United States of America and ITS-5G in Europe, vehicle-to-everything (V2X) over long-term evolution, and 5G cellular networks [2]. These connected vehicle technologies have been considered an integral part of modern Intelligent Transportation Systems (ITS). It can help enable various applications and services, from optimizing road traffic efficiency to improving road safety, from autonomous and connected vehicles to infotainment. These next-generation systems will improve the lifestyle of billions of people across the globe and influence society.

One interesting and promising application of using connected vehicle technologies for ITS is cooperative vehicle platoons. In the case of vehicle platooning, one vehicle escorts a group of vehicles. These vehicles operate on a highway or expressway, one behind another with small gaps in between. Generally, the leading vehicle is in manual driving mode and all other vehicles are set in automatic driving mode. They get the driving instructions from the leading vehicle. All the participating vehicles establish a network and disseminate information among themselves. Figure 1 illustrates the environment for vehicle platoons. There can be different variants of vehicle platoons. A homogeneous vehicle platoon involves a single type of vehicle and a heterogeneous vehicle platoon consists of different kinds of vehicles. Besides, a platoon of cargo vehicles can have all the vehicles owned by one company destined for a particular location. Maneuverings become a tad easier in this type of platoon. A different kind of platoon can be formed on the fly by a group of vehicles meeting for the first time on the road. A vehicle can advertise to offer its services as a leader over a wireless network. Other vehicles may join or leave the platoon depending on their source and destination addresses. A high-density platoon should have smaller gaps between vehicles while operating. These inter-vehicle distances are expected to be near about one-meter mark [3].

When vehicles form a platoon and operate with small gaps, the overall aerodynamic drag will be lesser. The vehicles have to apply more inferior force against the air while operating. Hence, the overall fuel consumption of the vehicles is reduced. It helps towards a greener environment and cheaper transportation. However, the amount of fuel saved may depend upon various factors like the types of vehicles, number of vehicles, wind directions, etc. Besides, more vehicles can fit into a small road segment with vehicle platoons. Thus, the highways or expressways could handle higher vehicle density leading to better utilization of the road infrastructure. Another advantage of vehicle platoons is the reduction in human efforts. Since only the leading vehicle requires manual driving, all other drivers can sit back comfortably or perform other work without worrying about driving. Even, there may be no need for drivers in the following vehicles in cargo vehicle platoons. Thus, the cargo trips would be more economical, and the saved human efforts could be utilized elsewhere. Moreover, computers are way faster than humans in processing and executing instructions compared to human beings. Thus, there is a huge chance that traffic control will become faster and safer, and the transit time of vehicles will also reduce.

II. Motivation and Background

In spite of having the ability to provide greener, more economical, and safer transportation, vehicle platooning has some striking issues inherited by the use of vehicular networks and connected vehicle technologies, which are generally absent in wireless networks. Vehicular platooning has stringent Quality-of-Service (QoS) requirements. For example, the 5G Automotive Association has identified latency and reliability as necessary QoS parameters for high-density platooning to exchange control messages in order to change speed or length, hand over to the driver, and terminate the platoon [4]. It becomes even more challenging to fulfill QoS requirements due to the dynamics of vehicle platoon environments, such as the ever-changing kinematics of vehicles, variable neighborhoods, fast and varying channel propagation, and interference from other communication systems in the vicinity. Besides, modern-day vehicles expected to participate in vehicle platooning will have many state-of-the-art sensors and devices such as cameras, lidar, and speedometers, alongside high-performance computing and sophisticated storage systems. They will generate, accumulate, process, store, and transmit a vast amount of data during a trip. This huge data provides plenty of opportunities with multiple dimensions for exploring the new designs of reliable and robust vehicular networks pertaining to vehicular platoons. Conventional networking strategies and algorithms are not suitable for handling such vast data and exploiting information for the betterment of the system. Therefore, it is necessary to think about alternative approaches.

Under the umbrella of artificial intelligence, machine learning significantly helps develop intelligent systems to operate in complex and hybrid environments. It has found success in different domains such as medical diagnostics, image and speech recognition, stock market prediction, and product recommendations. It involves the analysis of a vast amount of data by finding specific patterns and underlying designs to develop efficient machine-learning methods. It makes machine learning a set of data-driven techniques that can robustly handle heterogeneous data since there are no particular assumptions about the distribution of data during the process. Machine learning, especially deep and reinforcement learning, provides a broad set of tools that has the potential to analyze, classify, and predict situations based on mining the huge size of available data generated during communications between vehicles of a platoon. This can keep the platoon system more informed about the communication environment and help make better decisions for channel selections and QoS provisioning. Nonetheless, it is challenging to adapt the deep and reinforcement learning tools and exploit them for effective decision-making in vehicle platoons. It is a promising open area for research. This article aims to investigate and shed light on this prominent research field since applied machine learning in vehicle platoons is in its inception. Besides, this article also proposes employing deep reinforcement learning for link sharing in vehicle platoons as a new and intelligent example application.

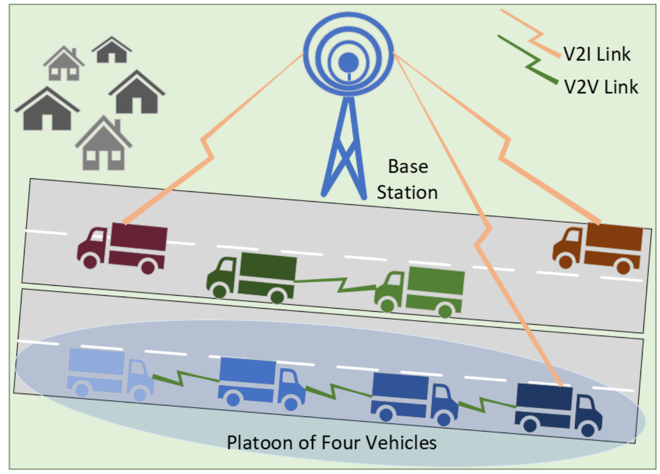

Fig. 1. Vehicle Platoon Environment: Heading out on an Expressway.

III. Machine Learning Implements for Vehicle Platoon Communications

Machine Learning implements can be broadly divided into Supervised, Unsupervised, and Reinforcement learning categories. However, different combinations of these learning implements could result in other learning implements such as online and semi-supervised. Table I illustrates the role of various machine learning techniques in solving issues pertaining to vehicle platoon communications. Machine learning works on the strategies of training and testing. In the training phase, the model learns with the provided data, and in testing the learned model is employed to produce the desired results.

A. Supervised Learning

In supervised learning, the model considers a labeled data set and tries to map the feature space to the decision space (i.e. labels). Depending upon discrete or continuous training labels the supervised learning is divided into classification and regression methods. In classification methods, incoming labels are categorized into various labels. Whereas, regression methods continuously predict the values for incoming samples. In communications and networks related to vehicle platoons, the classification methods can help in finding out the anomalies in the system, detecting the intrusions, and selecting the base stations [5]. On the other hand, regression methods can help in predicting parameters like delay and throughput related to the communication channels [7].

B. Unsupervised Learning

It is not always possible to have labeled data sets in real-world scenarios, and in such cases working with data is not easy. The answer to the problem is unsupervised learning methods. This type of method identifies structures and variables hidden in the data with the help of Bayesian learning to find the well-organized delineation of sample data. Clustering is the most used method of unsupervised learning where data points are clubbed together based upon some similarity criteria

TABLE I
MACHINE LEARNING FOR VEHICLE PLATOONS

Type	Function	Implements	Application Examples	Benefits	References
Supervised learning	Classification	Support vector machine, Neural networks, K-nearest neighbor	Detection of intrusion, fault, and anomaly, Vertical Control, Base station allocation	Relations between parameter, local decision making, flexibility	[5], [6]
Supervised learning	Regression	Support vector regression, Logistic regression	Delay prediction, Throughput prediction	Better vehicle's parameter prediction, robust, adaptive	[7], [8]
Unsupervised learning	Clustering	K-means	Driver assistance, Congestion control	Improved decision-making, better coordination, scalability	[9], [10]
Unsupervised learning	Reduction	Isometric Reduction, Manifold learning	Data aggregation over v2v communications	Real time adaptation, enhanced coordination and control	[11]
Reinforcement Learning	Policies	Q-Learning	Efficient message deliveries for platoon formation and maintenance	Improved communication efficiency, decentralized, better control	[12], [13]

to form the clusters. For vehicle platooning, it can be used for sending messages over communication links to assist the drivers and control congestion [10]. Another unsupervised learning method is reduction, where the high-dimension data is reduced to lower-dimension data without dropping much information since it is hard to deal with high-dimension data many a time. Isometric reduction and manifold learning are useful implementations for data aggregation and dimension reduction. The vehicle platoon captain can employ these data reduction techniques to reduce and aggregate the data before sending it to the base station and other members of the platoon over the communication channels [11]. It shall reduce the overall communication costs.

C. Reinforcement Learning

In reinforcement learning methods, an agent with no explicit supervision capabilities reads about environmental conditions by interacting with them. The agent maps the current environment condition with an action strategy to optimize the rewards. It uses the trial-and-error searching approach to achieve this goal. Usually, reinforcement learning considers Markov random process with actions and delayed rewards as a model for decision-making. Q-function is generally employed to solve the Markov random process-based problems. It is a model-free classical learning method and requires no additional information from the environment. It evaluates the expectation value for aggregate reward while choosing an action in a state. The optimal value of the Q-function is the maximum aggregate reward among all available aggregate rewards. The corresponding action to the optimal Q-function is selected as the strategy against the current environment scenario. Reinforcement learning implements can be used to solve the temporal and other network-parametric issues that arise during the control message exchanges during vehicle platoon formation and maintenance [12], [13]. However, reinforcement learning can be better applied to solve the vast range of problems related to communications in vehicle platoons in combination with deep learning.

IV. DRL FOR VEHICLE PLATOONS

Deep learning is a variant of an artificial neural network that is constructed with multiple layers of neurons. Recently, it has gained a lot of attention in solving problems in different fields. It focuses on learning the pattern and structures of the data, which can be provided as an input of any of the three main machine learning classes. In Deep Neural Networks (DNN), there is an input layer in which each neural node contributes toward a dimension of the data. An output layer is a combination of neural nodes responsible for providing desired output at the other end of the artificial neural networks. In between the input and out layers, there are several hidden layers. An artificial neuron represents a nonlinear mathematical transformation. It can be a sigmoid function, weighted sum, or unit function. The number of hidden layers and the number of nodes dictate the performance of the deep neural network. As the network goes deeper, its ability to represent data grows. However, with the increase in the number of hidden layers, challenges also increase, such as gradient exploding or vanishing. Also, it requires a bigger size of data for training [14]. However, with swift resources and better training mechanisms, these deep neural networks are a reality now. Deep learning has been successfully employed in various research domains like natural language processing, image processing, and computer vision, and the outcomes have refined the state-of-the-art in these domains. It is one of the motivations that dragged researchers into trying and testing it in their research domains. The results are encouraging, especially when used with machine learning methods, i.e., supervised, unsupervised, or reinforcement learning.

Deep learning studies a large amount of data and learns about the structure hidden in the data or discovers a pattern. However, after discovering the data patterns, it cannot make decisions on its own to solve the problems. Sometimes, reinforcement learning handles the decision-making part. Thus, the approach is called Deep Reinforcement Learning (DRL). The complex real-world problems where the environment is not structured and has a vast amount of data to be analyzed for decision-making are within reach of DRL. It has been evolving rapidly day by day with wide-ranging applications from gaming, robotics, healthcare, smart cities, education systems, and business activities to autonomous vehicles [14].

Vehicle platoon can benefit from the state-of-the-art DRL techniques. However, only a few works have been reported in the literature for vehicle platoon communications per se. Some of the related research domains and adopted DRL approaches are presented below. Also, the potential areas of research in the domain of vehicle platoons where DRL can be employed to integrate it with ITS of the future are discussed.

A. DRL for Platoon Control Strategy

Platoon formation and coordination require the exchange of control messages between vehicles for smooth operations. There are constraints and uncertainty in controlling the vehicles, that emerge out of the dynamic environment while operating as a platoon. The sequence of stochastic decisions needs to be made for it. DRL is used to deal with these problems in controlling the vehicles. DRL intelligently selects the external information of use thereby creating a better search space. Now, a better optimal strategy can be selected to reduce uncertainty and increase performance by deducing the effect of dynamic situations [15].

B. DRL for Traffic Control Strategy

The multi-agent DRL techniques are now being adopted in vehicle platoons to solve some different problems such as for adaptive cruise control. On the other hand, DRL techniques are also becoming much more useful in maintaining the large-scale traffic volume for future ITS [16]. Thus, it would be much easier to integrate the vehicle platoons with DRL strategies into the ITS. This should lead to seamless communication and exchange of control messages of DRL agents running on different entities of the unified system.

C. DRL for Parametric Optimization in Vehicle Platoons

Vehicle platoons are envisioned as the future of ITS, specifically with autonomous vehicles. However, due to increased traffic on roads and complex networks of roads, it becomes very crucial to optimize various parameters associated with vehicle platoons, such as fuel and energy consumption, transit delays, paths, etc. The traditional schemes require high computation powers, which are generally not available in the vehicle. DRL-based approaches could be handy in solving such problems. The non-linear relations between various input and output parameters could be modeled and optimized with the help of DRL techniques [17].

D. DRL for Mixed Vehicle Platoon Networks

In modern vehicle platoon scenarios, vehicles could be human-driven or could be autonomous, or could be a mix of both. Also, vehicles could be homogeneous or heterogeneous types. To add to this, there could be a scenario where multiple vehicle platoons are formed, which can communicate with each other for different purposes. To handle all these situations, there are works in the literature that show that DRL strategies can be useful. The Multi Actor Attention Critic strategy can deal with the spatial relationship of vehicles and can be an effective control strategy in mixed platoon scenarios [18].

V. DRL based Intelligent Resource Allocations

There are multiple exciting research works on resource allocation for device-to-device communications. However, these existing works are not for vehicle platoons and can not be employed directly for vehicle platoons because of their specificity. Most of the available are centralized and have a centralized controller that takes all the decisions for all the participating devices in the system. This kind of arrangement does not fit well for communication resource allocations in vehicle platoons because it incurs considerable overheads. Moreover, distributed resource allocation techniques used in other wireless network systems are not directly applicable to vehicle platoons due to the variation in surrounding environments [14]. Vehicle platoon applications with stringent

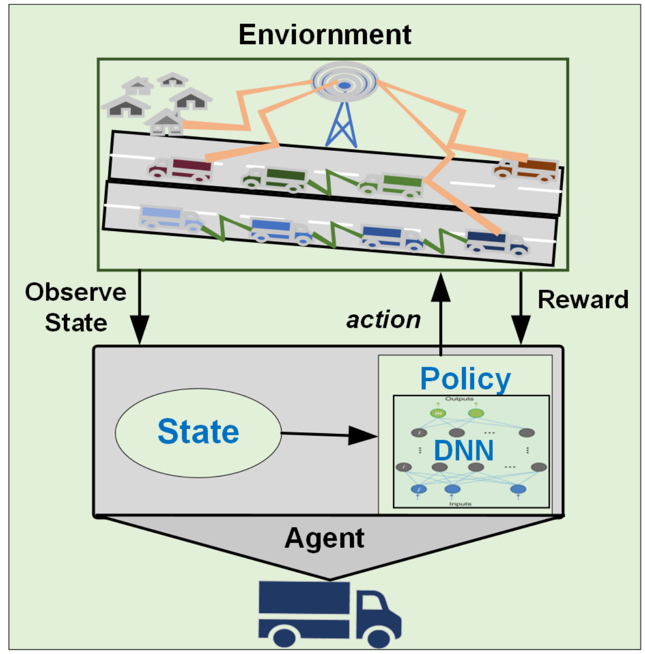

Fig. 2. DRL for Intelligent Resource Allocation in Vehicle Platoons

QoS requirements such as latency, reliability, etc., can be satisfied if the resources are allocated intelligently, considering local and temporal factors such as the kinematic of vehicles, inter-vehicle distance, and interference. Therefore, a new deep reinforcement learning-based intelligent and distributed communication resource allocation strategy is developed.

The main objective of the proposed method for resource allocation i.e., power and channel, is to meet the delay constraint for every vehicle platoon communication link i.e., vehicle-to-vehicle (V2V) links and reduce the obtrusion with vehicle-to-infrastructure (V2I) links. It is assumed that the V2I links are assigned with orthogonal channels in advance. The methodology for the proposed DRL scheme is illustrated in Fig. 2. Here, the agent, environment, perceptions, and actions are mentioned. We consider that an agent associated with a V2V link precept from the environment, which is everything apart from the V2V link. Since other V2V links can be rational, their actions such as the selection of resources, should be assumed as part of the environment. Depending upon the perception of the environment an agent takes appropriate action to select the power and channel for communication.

To model the problem, we assume that there is a state space $\mathcal{S}$, from which a V2V link, working as an agent, observes the current state s_τ of the environment at time τ. Based on that, the agent selects an action a_τ from the available set of actions, i.e., $\mathcal{A}$ using the decision policy p. It corresponds to designating a transmission power and communication channel. It is to transmit the remaining load, $\mathcal{M}_\tau$ by keeping the remaining transmission time in mind, such that the delay constraints, $\mathcal{D}_\tau$; are satisfied. Thus, the policy, p; can be determined by the Reinforcement Learning using the Q function, i.e.,

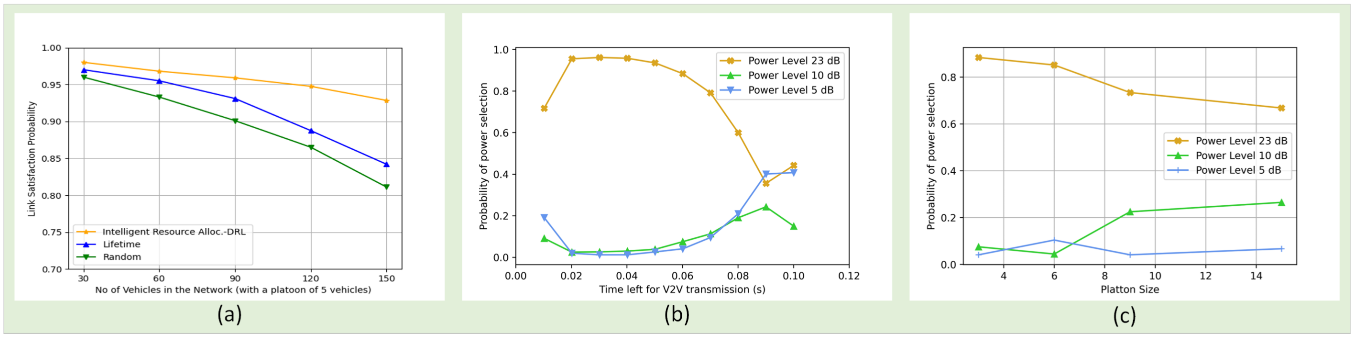

Fig. 3. (a) Delay satisfaction probability for V2V links in the vehicle platoon of size 5 vs. number of vehicles on road, (b) Power selection probabilities vs. time required to transmit the data, (c) Power selection probabilities vs. the platoon size

$Q(s_\tau, a_\tau, \theta)$ that maximizes the reward $\mathcal{R}_\tau$; of the agent. Here, θ is the parameter of the Q function. The performance of the Q function depends upon the value of θ; since:

$$\theta^* = \underset{\theta}{argmax} \mathcal{R}(\theta) \tag{1}$$

Here, the Deep Neural Network can determine its optimal value, θ^*; to narrow down to the best policy.

In this model, as the consequence of an action, the environment state changes from s_τ to $s_{\tau+\delta\tau}$. The state information parameters obtained by an agent to characterize the surrounding environment are the current channel information of the V2V link, i_τ; the present channel information of the V2I link, j_τ; the channel selected in the previous round, $\mathcal{C}_{\tau-\delta\tau}$; the channels selected by neighboring vehicles, $\mathcal{C'}_{\tau-\delta\tau}$; and the interference observed at an earlier state $\mathcal{I}_{\tau-\delta\tau}$. With the help of these parameters, the environment state at time τ can be determined by $s_\tau = [\mathcal{M}_\tau, \mathrm{i}_\tau, \mathrm{j}_\tau, \mathcal{C}_{\tau-\delta\tau}, \mathcal{C'}_{\tau-\delta\tau}, \mathcal{I}_{\tau-\delta\tau}]$. To create the environment, we considered the 3GPP [19] for the channel model (line-of-sight and none-line-of-sight) and highway scenarios (with six lanes, three in each direction).

The data generated using the 3GPP simulation model is stored in the memory to train and test our method. The Q learning algorithm, deep networks, and experience replays are used in training it. The sample data is taken from the memory for grounding the deep network, which reduces the temporal dependencies. The training starts with a random selection of policies which subsequently converges toward an improved policy as the model updates itself.

In fig. 3 (a), we depict a comparison of DRL-based resource allocation with random resource allocation and lifetime allocation schemes [20]. The DRL-based scheme has a higher probability of satisfying the constraint (delay) for V2V links in the vehicle platoons. This is because the proposed method can dynamically regulate the power levels and subbands to transmit the messages. Figure 3 (b) shows that when there is an ample amount of time left for transmission, the probability of selecting a lower power level increases. Whereas, when the remaining transmission time is less, the probability of selecting a higher power increases. Figure 3 (c) shows that the strategy continues even if the platoon size increases. This demonstrates some level of intelligence by agents in resource allocation using the proposed DRL scheme, which is absent in other existing strategies such as random and lifetime allocations.

VI. Real-Life Issues and Future Outlook

Vehicle platoon communication properties are distinguished, and merely applying DRL is inadequate. Thus, to deal with such properties, making adjustments to existing DRL implements or designing them for vehicle platoon specifically, is a challenging issue. Some of the critical requirements are mentioned below.

A. Distinguish Features of Vehicle Platoons

Vehicle platoon and its communication system both of them are dynamic. The traffic conditions, wireless channel propagation, kinematics of the platoon, and ever-changing neighborhoods play an essential part in creating such dynamics. Predicting such dynamics with historical data so that the correct action can be taken is still challenging. The advent of deep network-based models brings opportunities for improving predictions leveraging on the identifying dependency in the long term. For example, deep neural network-based prediction for communication channel allocations can be improved by using received signal strength and available data. However, it has yet to be studied whether deep neural networks can replace the traditional methods or not. Therefore, more endurance is required to dig deep in this research area.

B. Crossing Over the Gap between Simulation and Realism

Simulations are the current practice for the design and analysis of DRL techniques in the case of vehicle platoons. Also, it is the source of extensive data that can be analyzed further. However, having the vehicle platoon testbed is a costly affair and is still in the development phase. Testing and fine-tuning the DRL techniques is a complex task in real-world scenarios. The automobile industries started testing vehicle platoons [21]. Moreover, with the help of on-board units mounted on the vehicle and with certain assumptions, the testing can be performed in a controlled environment. [22].

Hopefully, the DRL techniques developed specifically for the vehicle platoons can be tested on the testbed.

C. DRL: A Heavy Weight Model

In vehicles, platoons can be homogeneous or heterogeneous. In both cases, vehicle platoon members have limited onboard processing resources. Moreover, it is not feasible to use remotely located resources such as clouds due to the restricted latency needs. On the other hand, DRL implementations are capable of extracting important and relevant information accurately with the help of available raw data but need high-performance computing resources. The solution to this issue is to enhance the computing resources of at least the leading vehicle, which is not a good idea. Another way is to design deep DRL implements specific to vehicular platoons, which could mitigate the resource requirement without hampering the performance of the system. In such cases, model compression and reduction methods could be helpful.

VII. Conclusion

In this article, we summarize the use of Deep and Reinforcement Learning techniques for vehicular platoons. It begins with an introduction to the fundamentals of the broad umbrella of artificial intelligence and it narrows down through machine learning and its categories, covering key features and notable algorithms for vehicle platoons. We then present some initial instances of applying DRL for decision-making within vehicle platoon scenarios. Next, we have proposed an intelligent resource allocation mechanism employing Q-learning and deep learning specific to vehicle platoons for effective communications. Finally, we shed light on research challenges warranting further investigation within the use of DRL for vehicle platoon communications.

References

[1] Oica, "2021 Statistics — www.oica.net — oica.net," https://www.oica.net/category/production-statistics/2021-statistics/, [Accessed 26-Apr-2023].

[2] L. Liang, H. Peng, G. Y. Li, and X. Shen, "Vehicular communications: A physical layer perspective," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 10 647–10 659, 2017.

[3] J. Hu, P. Bhowmick, F. Arvin, A. Lanzon, and B. Lennox, "Cooperative control of heterogeneous connected vehicle platoons: An adaptive leader-following approach," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 977–984, 2020.

[4] 5GAA, "5GAA Releases White Paper on C-V2X Use Cases: Methodology, Examples and Service Level Requirements - 5GAA — 5gaa.org," https://5gaa.org, [Accessed 27-Apr-2023].

[5] C.-C. Ho, B.-H. Huang, M.-T. Wu, and T.-Y. Wu, "Optimized base station allocation for platooning vehicles underway by using deep learning algorithm based on 5g-v2x," in *2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)*. IEEE, 2019, pp. 1–2.

[6] S. Boddupalli, A. S. Rao, and S. Ray, "Resilient cooperative adaptive cruise control for autonomous vehicles using machine learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 15 655–15 672, 2022.

[7] C. Yuan, Y. Li, H. Huang, S. Wang, Z. Sun, and H. Wang, "Application of explainable machine learning for real-time safety analysis toward a connected vehicle environment," *Accident Analysis & Prevention*, vol. 171, p. 106681, 2022.

[8] F. Jaffar, T. Farid, M. Sajid, Y. Ayaz, and M. J. Khan, "Prediction of drag force on vehicles in a platoon configuration using machine learning," *IEEE Access*, vol. 8, pp. 201 823–201 834, 2020.

[9] H. Borhan, M. Lammert, K. Kelly, C. Zhang, N. Brady, Y. Chia-Siung, and J. Liu, "Advancing platooning with adas control integration and assessment test results," *SAE International Journal of Advances and Current Practices in Mobility*, vol. 3, no. 0429, pp. 1969–1975, 2021.

[10] C. Chen, Y. Zhang, M. R. Khosravi, Q. Pei, and S. Wan, "An intelligent platooning algorithm for sustainable transportation systems in smart cities," *IEEE Sensors Journal*, vol. 21, no. 14, pp. 15 437–15 447, 2020.

[11] Y. Lou, P. Li, and X. Hong, "A distributed framework for network-wide traffic monitoring and platoon information aggregation using v2v communications," *Transportation research part C: emerging technologies*, vol. 69, pp. 356–374, 2016.

[12] M. F. Ozkan and Y. Ma, "Distributed stochastic model predictive control for human-leading heavy-duty truck platoon," *IEEE Trans. on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 16 059–16 071, 2022.

[13] A. Berbar, A. Gastli, N. Meskin, M. A. Al-Hitmi, J. Ghommam, M. Mesbah, and F. Mnif, "Reinforcement learning-based control of signalized intersections having platoons," *IEEE Access*, vol. 10, pp. 17 683–17 696, 2022.

[14] H. Ye, L. Liang, G. Y. Li, J. Kim, L. Lu, and M. Wu, "Machine learning for vehicular networks: Recent advances and application examples," *ieee vehicular technology magazine*, vol. 13, no. 2, pp. 94–101, 2018.

[15] L. Lei, T. Liu, K. Zheng, and L. Hanzo, "Deep reinforcement learning aided platoon control relying on v2x information," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 5811–5826, 2022.

[16] T. Chu, J. Wang, L. Codecà, and Z. Li, "Multi-agent deep reinforcement learning for large-scale traffic signal control," *IEEE transactions on intelligent transportation systems*, vol. 21, no. 3, pp. 1086–1095, 2019.

[17] C.-C. Yen, H. Gao, and M. Zhang, "Deep reinforcement learning based platooning control for travel delay and fuel optimization," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2022, pp. 737–742.

[18] Y. Xu, Y. Shi, X. Tong, S. Chen, and Y. Ge, "A multi-agent reinforcement learning based control method for connected and autonomous vehicles in a mixed platoon," *IEEE Transactions on Vehicular Technology*, 2024.

[19] 3GPP, "3rd generation partnership project: Technical specification group radio access network: Study lte-based v2x services: (release 14) standard 3GPP TR," https://portal.3gpp.org/, [Accessed 28-Nov-2022].

[20] N. S. Rajput, S. K. Satapathy, and S. Sisodia, "On stochastic modeling of robustness parameter for resilience communications in vehicle platoons," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 2022, pp. 946–951.

[21] "Automated platooning – step by step — scania.com," https://www.scania.com/group/en/home/newsroom/news/2018/automated-platooning-step-by-step.html, [Accessed 29-02-2024].

[22] N. S. Rajput, "Measurement of ieee 802.11 p performance for basic safety messages in vehicular communications," in *2018 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*. IEEE, 2018, pp. 1–4.