# Energy optimization in autonomous driving using deep reinforced learning and stochastic system modelling

[1]Akash Ladha, [2]Divya Priyadharshini Mohan

[1]Engineer, [2]Assistant Manager
[1]University of Michigan, [2]Lakshmi Machine Works

***Abstract*: In the development of future low-emission vehicles, Machine Learning is playing a significant role, since manufacturers progressively hit constraints with existing technology. In addition to independent driving, new improvements in reinforcement learning are also quite good at handling complicated parameterisation challenges. Deep reinforced training is utilised in this research to derive efficient electric hybrid vehicle operating methods. A wide range of possible driving and traffic conditions should be predicted, so that fuel-efficient solutions may be achieved in order to achieve intelligent and adaptable processes. This study demonstrates a reinforced learning agent's capacity to learn almost optimum operational strategies without previous route knowledge and gives a large potential for more factors to be included in the optimization procedure. This paper includes (1) a deep learning context that will enable discovering virtually optimum operating strategies. (2) The use of stochastic driver models to increase public generalisation and prevent overfitting of the approach. (3) Inclusion of the optimization process of battery modelling with extra power restrictions. The results are simulated and comparison graphs are plotted for the derived model.**

***Keywords*: Electric vehicles, Fuel economy, Exhaust emissions, Battery state of charge, Deep reinforced learning**

## I. Introduction

With the rising climate crisis, the demand for sustainable alternatives has seen a spike in recent years. Added to it, the exponential rise in population in countries like India is taking a toll on the lives of children [1]. Global Burden of Disease 2017 has reported that every three minutes, a child dies due to air pollution-related problems. State of Global Air 2020 reported that 116,000 infants in India died within a month of birth in 2019 due to air pollution [2]. Not alone in infants, air pollution has its evil wings spread on the adults too. In the year 2019, nearly 1.7 million people died due to air pollution. Even under strict lockdown conditions in India, in the year 2020, nearly 1, 20,000 people died. At least 12000 people died in Bengaluru alone and so is the case in many major cities [3].

Transportation sources are the major cause of air pollution in India. Sources say that by the year 2030 there will be 200 million vehicles on road. Besides, the majority of air pollution-related deaths are linked to diesel vehicles and 66% of air pollution is due to diesel vehicles [4]. To address all these issues the government of India (GOI) has shortlisted 102 cities as highly polluted and working hard to reduce particulate pollution by 20 to 30 per cent by 2024 [5].

India, which accounts for around 6% of global CO2 emissions from combustion fuel, is the third-largest nation worldwide to emit carbon. In 2019, 21 of the world's 30 most affected cities are in India, according to a survey released by IQAir[6]. Furthermore, 14 out of twenty the most polluted cities in the world are in India, according to the 2018 World Health Organization (WHO) GPD Database. The Indian Government is thus relying more and more on offering discounts and tax benefit incentives to afford high-cost electric rickshaws and to replace traditional partners [7].

Niti Aayog estimates suggest the EVs would end up saving up to 64 per cent of India's energy costs for road transport and cut down on up to 37 per cent of carbon emissions. The Indian government has already chalked down its roadmap with a promise of 30 per cent electric vehicles on the road by 2025. But the main challenge for Indian electric vehicle makers and the mobility ecosystem is that electric cars are affordable.

A significant number of local and small companies, who made up about two-thirds of the overall revenue of 2019 are dominating the Indian electric rickshaw industry primarily[8]. With few existing actors in the industry, the market has boomed over the past few years. YC Electric Vehicles is the industry leader in India in an electric rickshaw, with an integrated share of approximately 10 per cent. Mahindra & Mahindra Limited is followed by approximately 5 per cent.

The demand for Indian electric vehicles is rising rapidly. The unorganised local players in the industry are a large proportion of these markets. These unorganised players import and assemble parts of the vehicle locally, including engines, battery control systems and axles [9]. In recent years, founded Indian and foreign manufacturing companies for original equipment (OEMs) have also demonstrated their interest in the market, particularly concerning final-course connectivity [10].

OEMs are spending more and more on cheap and reliable electric rickshaws and well-known automotive players are launching new rickshaw models on this market. Due to their increasing business interest and ability to invest in product growth and distribution network expansion, the players will undoubtedly join the market in the coming years [11]. Some of the other key players operating in the Indian electric rickshaw market are Hero Electric Vehicles Pvt. Ltd., Terra Motors Corporation, Thukral Electric Bikes, ATUL Auto Ltd., Lohia Auto Industries, Kinetic Green Energy & Power Solutions Ltd., Electrotherm (India) Ltd., and Saera Electric Auto Pvt. Ltd.

Autonomous driving, with commitments to safety, convenience and energy economy, has been a priority for business and academics in the last decades [12]. The autonomous vehicle problems include other road users' unknown intents, communication among cars and with road infrastructure is a feasible technique for raising consciousness and enabling collaboration. Important research is happening into autonomous cars, in Europe and in other regions of the world, as well as in the USA and China [13]. It is just time for some persons in this field to outsource everyday travel to a computer by self-employed car before such improvements. In the near future urban autonomous transport networks, may be fitted with communication equipment and Global Positioning System (GPS), but with costly sensor technology[14].

## II. Literature Review

S Bacha et. al studied the control techniques for autonomous vehicle path following. In recent years, autonomous trajectories of electric vehicles have been significantly improved, especially as new sophisticated control algorithms have appeared. The major emphasis of these control algorithms is to ensure excellent performance with minimum distance errors for stability and tracking. This research presents an overview of the most frequently utilised techniques in autonomous electric vehicles. First of all, a geometric route tracking description is given which depends on the relationship between the dimension of the vehicle and its location on the path [15]. A kinematic path tracking is then explored to examine the reaction of the vehicle with regard to its speed and acceleration. Failure to use the forces on the vehicle may create some robustness concerns. Consequently, it is necessary to investigate dynamic route tracking with sophisticated control techniques.
The literature has well documented steering control for tracking in autonomous vehicles. In addition, continuous direct time control, e.g. vector control, is widely explored on human-driven electric vehicles with numerous engines. The combination of the two controls, however, is still not fully understood.
Chatzikomis et. al evaluates the advantages of torque vectoring in a self-sufficient electric vehicle, either by incorporating the torque vector system into the path tracking controller or by using it separately alongside the path tracking control. In obstacle avoidance experiments simulated with an empirically verified model for vehicle dynamics, a selection of tracking controllers is compared. For selecting controller settings, genetic optimization is utilised. The findings of the simulation show that torque vectoring is good for the autonomous reaction of the car [16]. When they are adjusted for the precise tire-road friction situation, the integrated controllers function best. However, while operating under reduced friction circumstances, they might potentially exhibit unstable behaviour without a refit. On the other hand, different torque vectors give a consistently steady cornering response to a variety of friction situations [17]. The route tracking performance is favourable for controllers with the Preview formulation or based on relevant reference pathways in relation to the mid-line lane.
Malan et al. discusses lateral dynamic electric vehicle control in an urban setting driven by public transit concerns to help minimise pollution in metropolitan regions. The framework under which the control strategies have been created is the "look-down reference," in which onboard sensors interact with road infrastructure to get lateral displacement. This mixed feedback structure is the intended control algorithm. The feedback action is carried out by three shackles, where the outside is nonlinear with cascade compensators. The feedback measurement is based on an awareness of the road curvature, which is rectified by the car's side displacement. The results of testing carried out on a real circuit indicate the effectiveness of the control technique discussed [18].
Guo et al examines the trajectory of autonomous vehicles with parametrically unsure, external disruptions and over actuated features after their control problem. The lateral movement of independent four-wheel drive electric cars is monitored by a new adaptive hierarchical control system. at first, an adaptive sliding mode high-level controlling law is designed for a frontal steering angle vector and a foreign yaw moment, in which a flexible logic technique adjusts the uncertain term and switching control gain to further moderate the chattering phenomenon and introduces an adaptive limiting layer [19]. Second, a pseudo inverse low-level control method is provided in order for the tire's longitudinal forces to be allocated best for the external yaw time. Finally, numerical simulations and practical results show the excellent tracking performance of the suggested adaptive control method. [20]
Although electric vehicles use green power and have green environmental benefits, the popularity of gasoline engine vehicles is still not gaining. The latest Electric Vehicle models feature distinct modes of driving assistance but are quite expensive. This is an autonomous system algorithm that changes and cheap costs, making electric vehicles smart and adaptable to use. The author has studied a four-wheel drive system and a double-axle rotation in our concept automobile to assure a fast reversal and short spatial rotation. The four-wheel drive system and double axel rotation are employed in your idea automobile and ensures fast U-turn and short rotation. They have utilised sophisticated self-controlling systems including automotive driving (limited to short distances), car parking, multiple mode driving aids, such as hill track detection, adaptive cruise control. This technology also detects unwanted rotation of the axle and can stop the automobile if impediments are found in front of it [21].
Past studies suggest that we need to have a dynamic and heuristic approach for the control of autonomous vehicles. We have concentrated in developing a dynamic control system for autonomous vehicles using deep reinforced learning in electric vehicles. The study also concentrates on energy saving optimization. For heterogeneous network electric vehicles (EVs), energy saving measures is extensively used [22]. The independent and precise regulation of motor- and regenerative braking torque may be performed separately and correctly on wheel motors including the in-wheel motor mounted in the wheel hub and close-to-the-wheel motor. Electric vehicles (EVs) for wheel motor (WMD) have recently been quickly developing. However, only a few studies carried out a thorough assessment of control techniques and uses of WMD EVs for energy efficiency.

## III. Methods

### 3.1 Reinforcement learning

Reinforcement learning is an agent who interacts with the environment, learns an optimum policy for sequential decision-making issues, both in natural sciences and social sciences as well as in engineering, via testing and error [23]. In the last years, in strengthening training, in the field of gaming, robotics, natural language processing, etc., deep learning or deep neural networks has prevailed [24]. Profound learning as a specific class is not without limits, for example as a blackbox that does not have interpretability, as a "alchemy" which is not subject to clear and adequate science rules, and which cannot compete with a child on tasks without human intellect [25].

Unsupervised learning tries, clustering and density estimation, to extract information from data without labels. The typical kind of unattended learning is representational learning. However, a kind of representational learning involves training feedback networks or convolutionary neural networks with supervised education[26]. Representation learning provides a way to maintain as much information as feasible about the original data, while keeping the display simpler and more accessible than the original data, in a low-dimensional sparing and autonomous way [27].
Deep neural networks automatically learn representations from raw inputs for recalling compositional hierarchies in various natural signals, i.e, lower level characteristics, e.g. pictures, objects hierarchy, components, motifs and local border combinations [28]. Distribution is a core notion in deep learning that means that every input may be represented by multiple features and that every feature may be numerous inputs. The deep, distributed representations' exponential benefits tackle exponential difficulties of the dimensions [29].
Mapping is one of the core driving pillars. Once an area is mapped it is possible to locate the present position of the car in the map. Google's initial reliable autonomous driving demos were based mainly on the location of pre-mapped places [30]. Due to the size of the challenge, the semantic object detection for accurate disambiguation increases standard mapping approaches. Also for object recognition, the localised high-development maps (HD maps) can be employed. An important aspect in the independent pipeline is trajectory planning. This module is needed to create motion-level instructions to guide the agent given a route-level plan using HD maps or GPS-based maps [31].
Classical motion planning ignores dynamics and differential restrictions while translating an agent into a target utilising translations and rotations [32]. A robotic agent which controls 6 degrees DOFs is considered holonomic, whereas an agent with less DOFs is known to be non-holonomic than its total DOFs. In the non-holonomical situation for autonomous drive classical algorithms, such as Djisktra algorithms, do not function. Random trees (RRT), which explores the configuration space quickly, are non-holonomic algorithms that explores the area with a random sample and free route and so we sue deep learning algorithm for motion control [33].
Motion planning is the task of ensuring the existence of a path between target and destination points. This is necessary to plan trajectories for vehicles over prior maps usually augmented with semantic information [34]. Path planning in dynamic environments and varying vehicle dynamics is a key problem in autonomous driving, for example negotiating right to pass through in an intersection merging into highways [35].

FIGURE 1INTERSECTION OF ROADS IN TRAFFIC

The system was trained in simulation first, and then trained in real time on board computers, and was able to train in a lane to complete a 250 m road test successfully. For learning models and policies directly from raw pixel inputs, model-based deep RL algorithms have been suggested [36]. Deep neural networks were utilised to produce predictions across hundreds of steps in simulated settings. Also for control, RL is useful. In comparison to RL techniques, classical optimum control methods such as LQR/iLQR are conducted. For the optimum control of stochastic situations, classical RL techniques are utilised [37].
Self-driving is an essential multi-agency work because, along with an agent's ego vehicle, many other actors such as pedestrians, bikers and other cars will also be present in simulated and real-world autonomous driving scenarios [38]. Therefore it is an essential future area of study to continue to build explicitly multi-agent learning methods to drive autonomous cars. Several previous techniques of autonomous driving tackled the problem with a MARL (Multi-agent reinforcement learning) viewpoint.

**3.2 Optimization of powertrain operation**

The existing EVs in the market often use power management methods heuristic or ad hoc which may not be optimal in a true pilot scenario. These techniques conservatively utilise the electricity source to prevent battery depletion [39. The optimisation of the powertrain is done by determination of the optimal real-time power split ratio. If the future trajectory of the vehicle speed is anticipated a priority, optimum power management may be accomplished [40].
Without connection data, this is unfeasible in real-time. This information enables the optimum controller to properly forecast the future pathways of the vehicle and to intelligently choose the right power source [41]. For instance, the controller may choose to use more electric energy when it is approaching a crossroads when it knows that the traffic signal turns red and a halt is essential. This reduces the operation of the engine and hence the usage of less gasoline while the battery is refilled when the car slows to a stop by regenerative braking [42].
Co-optimization of both car and engine means that connectivity is necessary to predict future traffic reports and partial or complete automation in order to implement the optimum vehicle control rule [43]. A three-step framework sample includes:

1. First, simulate the dynamics and dynamics of the longitudinal vehicle. The goal of optimization is to reduce the use of gasoline. The factors for control optimisation include the intended acceleration of the vehicle, motor throttle, gear shift and hybrid power use (if it exists). The information about the traffic (signal phase and time (SPaT)), the speed and acceleration of previous cars, the dynamic speed limit, etc.) are provided.
2. Secondly, identify the problem of energy efficiency optimisation and establish restrictions according to the individual CVAs. In Eco-AD, for example, the vehicle should keep its track safe, keep to its speed limits, and only pass when the light is green.
3. Third, create solutions to the aforementioned optimisation challenge that can be implemented in real time. The potential techniques include both the indirect method based on the calculus of variance and the minimal PMP (Particular Way) and

the direct method based on discretizing and resolving the problem by applying numerical optimization tools into a nonlinear programming (NLP) problem. [44]

Energy optimization includes the battery management, gear shift, power control. Using deep reinforcement learning, we try and optimize the battery aging, gear shift and power control [45].

### 3.3 System Modelling

We are trying to minimize the integral energy consumption value $E_{ev}$ for a time t, $t_0 \le t \le t_{des}$ for a minimum of battery usage $\dot{m}_{bat}$,

$$E_{EV} = \int_{T_0}^{T_{DES}} \dot{M}_{BAT}(ACTION(T), T)DT \qquad 1$$

Having the state of charge as boundary condition, where $SOC(t_{des})$ is the fully charged condition [46]. The power in the vehicle in influenced by the rolling friction and aerodynamic drag,

$$P_{VEHICLE} = F(V_{VEHICLE,}\Delta_{VEHICLE,}PAERO_{VEHICLE}) \qquad 2$$

The torque and speed of the vehicle is given as follows:

$$T_{ROLLING} = \frac{P_{VEHICLE}.R}{V_{VEHICLE}} \qquad 3$$

$$H_{VEHICLE} = \frac{V_{VEHICLE}}{2\Pi.R} \qquad 4$$

Where, r is the radius of the wheel [47].

With the speed and torque at hand, if we calculate the fuel consumed, then

$$m_{fuelcons} = f(\eta_{vehicle}, T_{vehicle}, G, P_{emel}) \qquad 5$$

To calculate the total power $P_{EV}$, we need to know the losses [48-51]. The following is the formula used in calculating the mechanical losses and electrical losses,

$$P_{mechanical} = f(\eta, p_{mechnical})$$
$$P_{electrical} = f(\eta, p_{electrical}) \qquad 6$$

battery is controlled as,

$$\vartheta_{battery,t+1} = f\left(\vartheta_{battery}, t\ , P_{battery}, Coolant_{Control}\right).....7$$

Where the state vector is defines as,

$$s_t = \left(n_{vehicle}, M_{vehicle}, SOC, \vartheta_{battery}, G\right) \qquad 8$$

Where G is the selected gear and SOC is defined as follows:

$$SOC_{t+1} = f\left(SOC_t\ , P_{battery}\right) \qquad 9$$

The power consumed by the electric vehicle is therefore given as,

$$P_{em} = \begin{cases} a_t, P_{em\max}\ fora_t \ge 0 \\ a_t, P_{em\min}\ fora_t \le 0 \end{cases} \qquad 10$$

Rewards are calculated as follows,

$$r_t = -\left(E_{bat} + \kappa_{Eelectrical}\right) \qquad 11$$

Where $E_{bat}$ Ebat and $\kappa_{Eelectrical}$ are calculated as follows,

$$E_{battery} = m_{battery} \cdot \rho_{battery} \cdot H_{battery} \qquad 12$$
$$E_{electrical} = P_{bat} \cdot \Delta t$$

The following diagram is as follows


FIGURE 2 : SIGNAL FLOW DIAGRAM OF ELECTRIC VEHICLE

The Q-learning algorithm is as follows [52-54]:
initialize Q arbitrarily, e.g., to 0 for all states, set action value for terminal states as 0
initialize the replay buffer
for each episode do initialize state s for each step of episode, state s is not terminal do
a ← action for s derived by Q, e.g., ε-greedy
choose action at with actor and add noise
execute at
observe rt+1 and st+1
store [st,at,rt+1,st+1] in buffer
choose minibatch from buffer
take action a, observe r, s’
Q(s, a) ← Q(s, a) + α[r + γ maxa0 Q(s, a0) − Q(s, a)]s ← s’
end
end

## IV. Results and Discussions

The results were simulated using open source software Scilab. The actor gets a standardised vector of the state and produces the stated continuous action. The critic also enters into the second network layer and produces a continuous Q-value through linear activation. A discount factor of $\gamma = 0$ was set for the training procedure. Therefore the agent optimised his approach exclusively for local rewards. However, given there is a dynamic weighting element based on the present charge status in the award function, the agent has adapted the technique quite far-sightedly. This means that considerably better outcomes than discounts $\gamma = 0$ were attained.

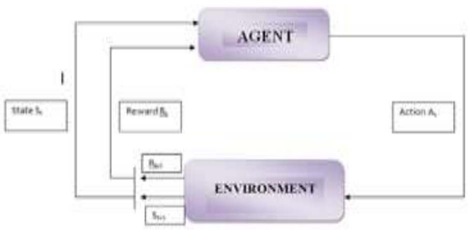


FIGURE 3 AGENT-ENVIRONMENTINTERACTION FOR REINFORCEMENT LEARNING

The engine speed was plot as such:
The resulting energy consumption is reduced according to the bonus function consistently, while the fuel saved by the combustion engine alone, which has been set very early, is more than 20%. With the resulting battery charge near to the starting value of 50 percent in the terminal state of an episode, SOC may be regarded as balanced. A positive variation of the cycle battery charge for any efficient operating strategy might easily be predicted in a low-speed situation where the low-powered electrical motor can be utilised extremely efficiently.

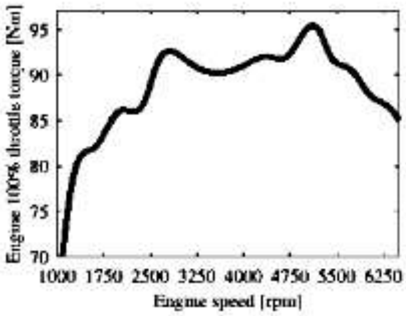


FIGURE 4 MOTOR SPEED VS TORQUE

The learning rate of the actor network had to be adjusted occasionally to driving profiles including rather high speed and accelerated rates, since the unavoidable increases in reward by increased momentary fuel consumption lead to greater degrees in the beginning of training, where neuronal networks are initialised with close to nil outputs and the loss increases acceleration It would be a substantial alternative to reduce the incentives or the slopes to specified maximum values, but important information regarding the use of fuel can be eliminated later in the training.

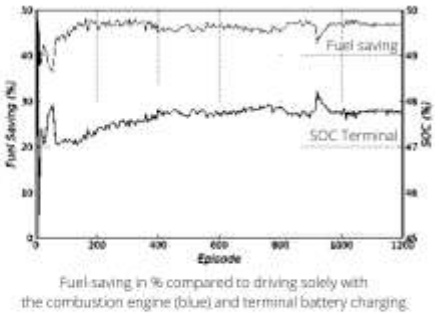


FIGURE 5 FUEL SAVING COMPARISION

In comparison to DP, the RL agent does not need past driving information and is taught to optimally manage the EM using just instantaneous vehicle environmental input. The benefit of deep reinforcement education above any other traditional approach for deriving HEV operational strategies is a local strategy optimization with almost worldwide optimal answers.

Stochastic cycles might also contain any kind of information on driving behaviours or traffic conditions that neural networks can simply consider in the operational plan. This gives a far more exact depiction of reality traffic which may result, rather than a generating approximate values on a test stand, in increasing fuel savings and reducing emissions in the actual operation of the vehicle.

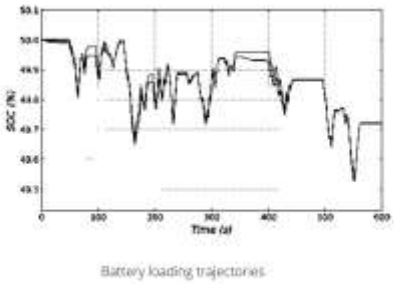


FIGURE 6 BATTERY LOADING TRAJECTORIES

## V. Conclusion

In real-world autonomous driving applications, strengthening learning is still an active and growing subject. Some successful commercial applications are available, there is relatively little public literature or large-scale dataset. We were encouraged to define and arrange autonomous driving RL applications. Autonomous driving situations include interactive agents and need to be negotiated and to be dynamically decided according to RL. Many problems have to be overcome to provide mature, detailed answers. Theoretical enhancement learning and an extensive literature analysis on RL application for autonomous driving challenges are provided throughout this study.

Developing explicitly multi-agent learning techniques to the problem of autonomous driving is also a significant and unheard of task for the future. MARL methods facilitate coordination and high-level decision making amongst autonomous vehicle groups and provide new chances for the safety of autonomous driving regulations to be tested and validated. In addition, RL algorithms are a difficult challenge for academics and practitioners. In conclusion, we believe this study will stimulate future research and applications.

**References**


[1] Zhang Y, Ai Z, Chen J, You T, Du C, Deng L. Energy-Saving Optimization and Control of Autonomous Electric Vehicles With Considering Multiconstraints. IEEE Trans Cybern. 2021 Apr 19;PP. doi: 10.1109/TCYB.2021.3069674. Epub ahead of print. PMID: 33872179.

[2] Fayyad J, Jaradat MA, Gruyer D, Najjaran H. Deep Learning Sensor Fusion for Autonomous Vehicle Perception and Localization: A Review. Sensors (Basel). 2020 Jul 29;20(15):4220. doi: 10.3390/s20154220. PMID: 32751275; PMCID: PMC7436174.

[3] Zhou X, Zhou J. Optimization of autonomous driving state control of low energy consumption pure electric agricultural vehicles based on environmental friendliness. Environ Sci Pollut Res Int. 2021 Apr 30. doi: 10.1007/s11356-021-14125-9. Epub ahead of print. PMID: 33928505.

[4] Gao K, Yan D, Yang F, Xie J, Liu L, Du R, Xiong N. Conditional Artificial Potential Field-Based Autonomous Vehicle Safety Control with Interference of Lane Changing in Mixed Traffic Scenario. Sensors (Basel). 2019 Sep 27;19(19):4199. doi: 10.3390/s19194199. PMID: 31569737; PMCID: PMC6806175.

[5] S. Bacha, M. Y. Ayad, R. Saadi, A. Aboubou, M. Bahri and M. Becherif, "Modeling and control technics for autonomous electric and hybrid vehicles path following," 2017 5th International Conference on Electrical Engineering - Boumerdes (ICEE-B), 2017, pp. 1-12, doi: 10.1109/ICEE-B.2017.8191998.

[6] C. Chatzikomis, A. Sorniotti, P. Gruber, M. Zanchetta, D. Willans and B. Balcombe, "Comparison of Path Tracking and Torque-Vectoring Controllers for Autonomous Electric Vehicles," in IEEE Transactions on Intelligent Vehicles, vol. 3, no. 4, pp. 559-570, Dec. 2018, doi: 10.1109/TIV.2018.2874529.

[7] S. Malan, M. Milanese, P. Borodani and A. Gallione, "Lateral Control of Autonomous Electric Cars for Relocation of Public Urban Mobility Fleet," in IEEE Transactions on Control Systems Technology, vol. 15, no. 3, pp. 590-598, May 2007, doi: 10.1109/TCST.2007.894657.

[8] J. Guo, Y. Luo and K. Li, "An Adaptive Hierarchical Trajectory Following Control Approach of Autonomous Four-Wheel Independent Drive Electric Vehicles," in IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 8, pp. 2482-2492, Aug. 2018, doi: 10.1109/TITS.2017.2749416.

[9] "Smart intelligent electric vehicle (Enhancing flexibility)," 2017 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2017, pp. 33-37, doi: 10.1109/ECACE.2017.7912874.

[10] Shao, Y., Mohd Zulkefli, M. A., and Sun, Z. (September 1, 2017). "Vehicle and Powertrain Optimization for Autonomous and Connected Vehicles." ASME. Mechanical Engineering. September 2017; 139(09): S19–S23. https://doi.org/10.1115/1.2017-Sep-6

[11] Holjevac N, Cheli F, Gobbi M. Multi-objective vehicle optimization: Comparison of combustion engine, hybrid and electric powertrains, Proc Inst Mech Eng, Part D: J Automobile Eng (2019) 0954407019860364.

[12] Zablocki, É., Ben-Younes, H., Pérez, P., & Cord, M. (2021). Explainability of vision-based autonomous driving systems: Review and challenges. arXiv preprint arXiv:2101.05307.

[13] Alawadhi, M., Almazrouie, J., Kamil, M., & Khalil, K. A. (2020). A systematic literature review of the factors influencing the adoption of autonomous driving. International Journal of System Assurance Engineering and Management, 11(6), 1065-1082.

[14] Kiran, B. R., Sobh, I., Talpaert, V., Mannion, P., Al Sallab, A. A., Yogamani, S., & Pérez, P. (2021). Deep reinforcement learning for autonomous driving: A survey. IEEE Transactions on Intelligent Transportation Systems.

[15] Cárdenas, J. F. S., Shin, J. G., & Kim, S. H. (2020). A few critical human factors for developing sustainable autonomous driving technology. Sustainability, 12(7), 3030.

[16] Chen, J., Li, S. E., & Tomizuka, M. (2021). Interpretable end-to-end urban autonomous driving with latent deep reinforcement learning. IEEE Transactions on Intelligent Transportation Systems.

[17] Graves, D., Nguyen, N. M., Hassanzadeh, K., & Jin, J. (2020). Learning predictive representations in autonomous driving to improve deep reinforcement learning. arXiv preprint arXiv:2006.15110.

[18] Kalweit, G., Huegle, M., Werling, M., & Boedecker, J. (2020). Interpretable Multi Time-scale Constraints in Model-free Deep Reinforcement Learning for Autonomous Driving. arXiv preprint arXiv:2003.09398.

[19] Li, G., Li, S., Li, S., Qin, Y., Cao, D., Qu, X., & Cheng, B. (2020). Deep reinforcement learning enabled decision-making for autonomous driving at intersections. Automotive Innovation, 3(4), 374-385.

[20] Aradi, S. (2020). Survey of deep reinforcement learning for motion planning of autonomous vehicles. IEEE Transactions on Intelligent Transportation Systems.

[21] Huegle, M., Kalweit, G., Werling, M., & Boedecker, J. (2020, May). Dynamic interaction-aware scene understanding for reinforcement learning in autonomous driving. In 2020 IEEE International Conference on Robotics and Automation (ICRA) (pp. 4329-4335). IEEE.

[22] Bhalla, S., Subramanian, S. G., & Crowley, M. (2020, May). Deep multi agent reinforcement learning for autonomous driving. In Canadian Conference on Artificial Intelligence (pp. 67-78). Springer, Cham.

[23] Almási, P., Moni, R., & Gyires-Tóth, B. (2020). Robust reinforcement learning-based autonomous driving agent for simulation and real world. arXiv preprint arXiv:2009.11212.

[24] Liu, X., Liu, Y., Chen, Y., & Hanzo, L. (2020). Enhancing the fuel-economy of V2I-assisted autonomous driving: A reinforcement learning approach. IEEE Transactions on Vehicular Technology, 69(8), 8329-8342.

[25] Revenga, P. A., Gutiérrez, R., & Dıaz, A. (2020). DQN-Based Deep Reinforcement Learning for Autonomous Driving. In Advances in Physical Agents II: Proceedings of the 21st International Workshop of Physical Agents (WAF 2020), November 19-20, 2020, Alcalá de Henares, Madrid, Spain (p. 60). Springer Nature.

[26] Chen, S., Wang, M., Song, W., Yang, Y., Li, Y., & Fu, M. (2020). Stabilization Approaches for Reinforcement Learning-Based End-to-End Autonomous Driving. IEEE Transactions on Vehicular Technology, 69(5), 4740-4750.
[27] Hoel, C. J., Wolff, K., & Laine, L. (2020). Tactical decision-making in autonomous driving by reinforcement learning with uncertainty estimation. In 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 1563-1569). IEEE.
[28] Zhou, M., Luo, J., Villela, J., Yang, Y., Rusu, D., Miao, J., ... & Wang, J. (2020). SMARTS: Scalable Multi-Agent Reinforcement Learning Training School for Autonomous Driving. arXiv preprint arXiv:2010.09776.
[29] Zhu, M., Wang, Y., Pu, Z., Hu, J., Wang, X., & Ke, R. (2020). Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving. Transportation Research Part C: Emerging Technologies, 117, 102662.
[30] Haydari, A., & Yilmaz, Y. (2020). Deep reinforcement learning for intelligent transportation systems: A survey. IEEE Transactions on Intelligent Transportation Systems.
[31] Ye, F., Cheng, X., Wang, P., Chan, C. Y., & Zhang, J. (2020, February). Automated lane change strategy using proximal policy optimization-based deep reinforcement learning. In 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 1746-1752). IEEE.
[32] Wu, Z., Sun, L., Zhan, W., Yang, C., & Tomizuka, M. (2020). Efficient sampling-based maximum entropy inverse reinforcement learning with application to autonomous driving. IEEE Robotics and Automation Letters, 5(4), 5355-5362.
[33] Park, S., Kang, Y., Tian, Y., & Kim, J. (2020, January). Fast and Reliable Offloading via Deep Reinforcement Learning for Mobile Edge Video Computing. In 2020 International Conference on Information Networking (ICOIN) (pp. 10-12). IEEE.
[34] Aoki, S., Higuchi, T., & Altintas, O. (2020, October). Cooperative perception with deep reinforcement learning for connected vehicles. In 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 328-334). IEEE.
[35] Benhamou, E., Saltiel, D., Ohana, J. J., & Atif, J. (2021, January). Detecting and adapting to crisis pattern with context based Deep Reinforcement Learning. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 10050-10057). IEEE.
[36] Osiński, B., Jakubowski, A., Zięcina, P., Miłoś, P., Galias, C., Homoceanu, S., & Michalewski, H. (2020, May). Simulation-based reinforcement learning for real-world autonomous driving. In 2020 IEEE International Conference on Robotics and Automation (ICRA) (pp. 6411-6418). IEEE.
[37] Fuchs, F., Song, Y., Kaufmann, E., Scaramuzza, D., & Dürr, P. (2021). Super-human performance in gran turismo sport using deep reinforcement learning. IEEE Robotics and Automation Letters, 6(3), 4257-4264.
[38] Grigorescu, S., Trasnea, B., Cocias, T., & Macesanu, G. (2020). A survey of deep learning techniques for autonomous driving. Journal of Field Robotics, 37(3), 362-386.
[39] Hart, P., & Knoll, A. (2020). Using counterfactual reasoning and reinforcement learning for decision-making in autonomous driving. arXiv preprint arXiv:2003.11919.
[40] Likmeta, A., Metelli, A. M., Tirinzoni, A., Giol, R., Restelli, M., & Romano, D. (2020). Combining reinforcement learning with rule-based controllers for transparent and general decision-making in autonomous driving. Robotics and Autonomous Systems, 131, 103568.
[41] Fernando, T., Denman, S., Sridharan, S., & Fookes, C. (2020). Deep Inverse Reinforcement Learning for Behavior Prediction in Autonomous Driving: Accurate Forecasts of Vehicle Motion. IEEE Signal Processing Magazine, 38(1), 87-96.
[42] Luo, Q., Li, C., Luan, T. H., & Shi, W. (2020). Collaborative data scheduling for vehicular edge computing via deep reinforcement learning. IEEE Internet of Things Journal, 7(10), 9637-9650.
[43] Xu, Z., Chen, J., & Tomizuka, M. (2020). Guided Policy Search Model-based Reinforcement Learning for Urban Autonomous Driving. arXiv preprint arXiv:2005.03076.
[44] Jang, I., Kim, H., Lee, D., Son, Y. S., & Kim, S. (2020). Knowledge Transfer for On-Device Deep Reinforcement Learning in Resource Constrained Edge Computing Systems. IEEE Access, 8, 146588-146597.
[45] Wu, Z., Qiu, K., & Gao, H. (2020). Driving policies of V2X autonomous vehicles based on reinforcement learning methods. IET Intelligent Transport Systems, 14(5), 331-337.
[46] Du, Z., Miao, Q., & Zong, C. (2020). Trajectory Planning for Automated Parking Systems Using Deep Reinforcement Learning. International Journal of Automotive Technology, 21(4), 881-887.
[47] de Morais, G. A., Marcos, L. B., Bueno, J. N. A., de Resende, N. F., Terra, M. H., & Grassi Jr, V. (2020). Vision-based robust control framework based on deep reinforcement learning applied to autonomous ground vehicles. Control Engineering Practice, 104, 104630.
[48] Amini, A., Gilitschenski, I., Phillips, J., Moseyko, J., Banerjee, R., Karaman, S., & Rus, D. (2020). Learning robust control policies for end-to-end autonomous driving from data-driven simulation. IEEE Robotics and Automation Letters, 5(2), 1143-1150.
[49] Benhamou, E., Saltiel, D., Ungari, S., & Mukhopadhyay, A. (2020). Bridging the gap between Markowitz planning and deep reinforcement learning. arXiv preprint arXiv:2010.09108.
[50] Chen, D., Jiang, L., Wang, Y., & Li, Z. (2020, July). Autonomous driving using safe reinforcement learning by incorporating a regret-based human lane-changing decision model. In 2020 American Control Conference (ACC) (pp. 4355-4361). IEEE.
[51] Fang, Wenqi, Shitian Zhang, Hui Huang, Shaobo Dang, Zhejun Huang, Wenfei Li, Zheng Wang, Tianfu Sun, and Huiyun Li. "Learn to Make Decision with Small Data for Autonomous Driving: Deep Gaussian Process and Feedback Control." Journal of Advanced Transportation 2020 (2020).

[52] Fu, Y., Li, C., Yu, F. R., Luan, T. H., & Zhang, Y. (2020). A decision-making strategy for vehicle autonomous braking in emergency via deep reinforcement learning. IEEE transactions on vehicular technology, 69(6), 5876-5888.
[53] Dong, J., Chen, S., Li, Y., Du, R., Steinfeld, A., & Labi, S. (2020). Facilitating connected autonomous vehicle operations using space-weighted information fusion and deep reinforcement learning based control. arXiv preprint arXiv:2009.14665.
[54] Lei, L., Tan, Y., Zheng, K., Liu, S., Zhang, K., & Shen, X. (2020). Deep reinforcement learning for autonomous internet of things: Model, applications and challenges. IEEE Communications Surveys & Tutorials, 22(3), 1722-176.