A Reinforcement Learning-Based Approach for Autonomous Vehicle Decision-Making

Bingquan Wang *

TEDA Maple Leaf High School, Tianjin, 300000, China

* Corresponding Author Email: 23030197@students.mapleleafedu.com

Abstract. Nowadays, cars are always the things people cannot live without. With the development and improvement of the car industry and general income, the number of cars equipped with self-driving systems is increasing dramatically in the market. Self-driving has already shown a tendency for a car to have this function as a standard instead of an attractive feature. In contrast, the reliability of this technology is now being doubted. Many drivers point out that sometimes their cars will make silly decisions while in self-driving mode, such as ignoring other cars on the road. This paper aims to come up with a possible approach to help in reducing the errors of the self-driving system in decision-making. The paper mainly uses a quantitative study through the literature review method in data collection and analysis. By summarizing data and reviewing existing materials, a new learning-based approach will be demonstrated in this paper. It discusses a feasible way to let AI in self-driving vehicles memorize and judge road conditions subjectively.

Keywords: Self-driving car, AI, Decision-making.

1. Introduction

This paper focuses on one specific concept of self-driving cars, which is decision-making. Now, many car brands have already installed this technology on their mainstream models. It brings tremendous benefits to reduce traffic stress and fatal accidents. For instance, Tesla was one of the pioneers of this field. All standard versions of its car types are equipped with a self-driving system. It is safe to say Tesla rewrote this technology from a stunt to practical. Autonomous driving can be generally divided into three major categories, which are perception, decision-making, and command. To be more specific, decision-making is the crucial part of all of them. It is similar to the human brain, dealing with various information and finally coming to a decision. Generally, there are three different ways of decision-making in autonomous driving. From simple to complex, decisions based on rules and reasoning, decisions based on optimization and search, and decisions based on learning. This paper emphasizes the last way to find a possible approach based on enhancing the learning ability of AI. Many experts and scholars have made contributions in this field. In contrast, their studies only focused on one specific road type, like a highway (Guan et al., 2018). It only solves issues that are demonstrated in one scenario. This technology still lacks an approach to make it comprehensive and easily adapts to various road conditions. The reinforcement learning (RL) approach perfectly fits AI in decision-making. It views the driving process as a sequential decision-making problem, where the vehicle is the agent that learns the optimal strategy through trial and error. It can explore decision-making strategies that surpass human ones and are more optimal, and be good at handling interactive games. There are many studies in other fields that have proved the success of the RL approach, like in bioimaging, medical imaging, and (brain/body)-machine interfaces (Mahmud et al., 2018).

It provides a paradigm for autonomous driving technology in decision-making. Admittedly, it still faces many limits and challenges. In other words, this approach has never been tested in the field of intelligent driving. To break it down, its dependability and safety guarantee still need to be tested. This paper will give a detailed approach that provides theoretical support. This RL approach is essential because it highlights AI to learn and analyze by itself. This pioneering plan offers unlimited possibilities and potential for the development of this industry. Generally, this paper will explain how AI is able to handle different driving scenarios through autonomous learning methods. It includes traffic jams, highway lane changing, and alleyways. These are selected to be representative driving

scenarios that drivers usually face. This might offer new ideas for car companies in the field of autonomous driving. What sets this research apart from previous ones is that it no longer only confines itself to the existing knowledge base of AI but enables AI to learn independently and sense road conditions to make decisions. More precisely, every time the car accelerates, decelerates, or turns, it will go through a series of learning processes to make judgments and execute accordingly. A new concept of “initiative brain” will be introduced as well. This provides a feasible technical path for automotive enterprises to develop safer and more reliable autonomous driving systems. This paper will demonstrate how the RL approach performs and applies in various driving scenarios. Besides, it will also come up with basic formulas to explain how this approach works.

2. Traffic Jam

To start with, the application of autonomous learning artificial intelligence (AI) in self-driving cars represents a paradigm shift from rule-based and supervised learning systems towards more adaptive, self-improving frameworks. Unlike pre-programmed responses or models trained solely on historical data, an AI equipped with enhanced autonomous learning capabilities utilizes a continuous loop of perception, decision-making, evaluation, and model optimization in real-time. This segment elucidates the practical application and profound significance of such a method within the specific context of heavy traffic jams, arguing that it is indispensable for achieving true Level 4 and 5 autonomies.

In a traffic jam scenario, the challenges are multifaceted and extend beyond simple vehicle following. Static navigation maps become unreliable, and the behavior of other drivers becomes increasingly erratic. A conventionally programmed autonomous vehicle might rely on conservative, rigid protocols, which, while safe, can be inefficient and disrupt the natural flow of traffic, even contributing to the propagation of "phantom traffic jams" (where a single deceleration causes a wave of stopping behind it) (Kesting & Treiber, 2013). An autonomously learning AI, however, approaches this environment as a rich source of live data for immediate learning and adaptation. Its application unfolds in a multi-layered process. The core of this process lies in the decision-making module, which is likely based on a deep reinforcement learning (RL) framework. To deeply and directly explain how this approach works, it can be expressed in a simple, basic formula. Assuming the information analyzed by the central processing unit (CPU) of AI marked as “C”, then it can be demonstrated as this.

$$\text{Decision} = \text{Road condition} * N + C \tag{1}$$

$$N = \text{Numbers of the road condition met before} \tag{2}$$

In this formula, the most crucial part is how the CPU analyzes the information given by sensors on a car. Generally, it is the most basic rule of this approach that works. In different scenarios, the formula will have some specific changes. In a traffic jam, there are always many possibilities that lead to congestion. Simply, it can be categorized into two major causes: contrived and natural. Rush hours, malfunctioning traffic lights, car accidents, and the leading vehicle moving too slowly can all be regarded as contrived factors. For those traffic jams caused by obstacles blocking the road, like collapsing trees or telegraph poles, or there is a construction that has to take up one of the lanes. All these factors can be regarded as natural factors. For AI, while autonomous driving is on, it struggles to drive smoothly in the flow of traffic. This driving scenario tests how flexibly the AI can weave through traffic and find the fastest route through congested sections. Now, most of the autonomous driving systems that people can find on their cars are limited to a protocol. When these versions of AI face traffic jams, they would be more likely to reactively brake instead of proactively motion (actively seek chances to cut in line). This strategy does indeed ensure the safety of drivers when

using the autonomous driving system during peak traffic hours, but it does not make any substantial changes to truly alleviate the inconvenience caused by traffic jams. Except for rush hour, most traffic jams are sudden in nature. They are generally caused by poor road traffic management, and they indirectly lead to massive waste of time, fuel, and money (Bhardwaj et al., 2023). Thus, the old typed breaking strategy can barely reduce this issue; instead, the RL approach perfectly adapts to this driving scenario. It highlights AI to be an active and flexible driver in the flow of traffic. It asks AI to design an ideal route that saves the least time and fuel for the driver. Next, this paper will give several formulas and images to explain how the RL approach works in this specific scenario.

$$\text{Decision} = \text{Road condition}*N+C \tag{3}$$

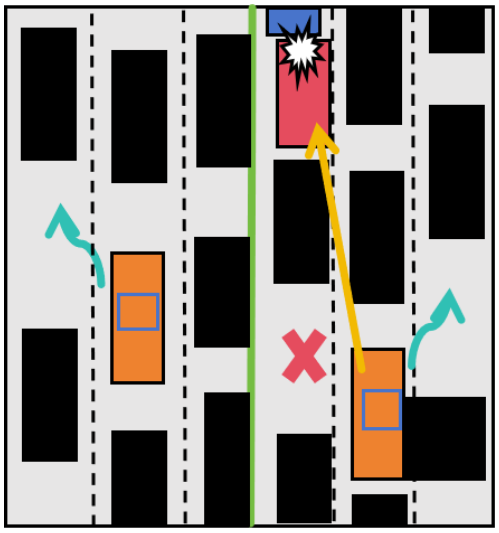


Figure 1. How the RL approach works in a traffic jam (a/b)

This is the original formula we had, now using images to simulate real traffic jam scenarios.

These are two common situations that a driver may face in a traffic jam. The orange rectangle is the car with an AI driving system. In Figure 1(a), the car is asked to shift the lane (cut in line) in a crowded traffic flow. To explain this process in detail, firstly, it will follow the basic formula to analyze the current driving status of the vehicle and calculate whether the distances between it and the other vehicles in front, behind, left, and right are within the safe passing distance. Then, the AI will actively seek the appropriate timing to make the decision to merge and complete the process smoothly. In Figure 1(b), the car will use one method in the RL approach, which is insight. It shows a certain degree of complexity in the logical thinking of AI. First, AI will analyze and calculate the driving environment as well as it did in (a). Second, the AI should identify which lane is available. In the image, there is a car accident that happened in the left driving lane. This RL approach allows AI to learn to use sensors and radars to subjectively form a clear vision not only confined to the surrounding environment but also at a far distance. For AI, there are two routes to choose from. It will select the right lane because this approach helps it obtain a result that the left lane is unavailable, while the right lane can be switched.

Admittedly, the actual traffic congestion involves factors that are more complex and diverse than those listed above. However, this RL approach focuses on enabling AI to autonomously sample and learn to respond to various driving conditions. Therefore, this approach contributes to alleviating traffic pressure by enhancing the autonomous decision-making ability of AI in driving.

3. Highway Lane Changing

This section will shift the focus to the most common scenario for autonomous driving - highways. Most people tend to use the autonomous driving system more frequently during long-distance highway driving than in other scenarios. In this field, there have already been related papers proposing

an RL learning based safety method that is applicable to vehicles in highway driving (Mohammadhasani et al., 2021). It proposes an algorithm for vehicles to avoid various situations they may encounter during driving (including avoiding collisions). In contrast, autonomous driving system still lacks the means to evade and handle unpredictable events (people, random obstacles). In different countries, the conditions of highways may vary to some extent. For instance, in America, most highways are mainly paved with asphalt. In urban areas, highways usually have some fences and buffer zones. Some interstate highways, however, are built entirely in the wild, almost in direct contact with nature. But overall, for vehicles, these are both ideal driving environments. In Malaysia, the quality of highways is far below the standards of the former ones (Rahman et al., 2020). Due to the shortage of financial resources and the complexity of the natural environment, the highways there usually have to be incorporated into rough paths or lack protective measures against sudden intrusions of pedestrians or even animals. These factors significantly increase the risks of using autonomous driving systems in certain districts of Malaysia.

In order to minimize the probability of accidents caused by unpredictable factors, the RL approach mentioned in this paper is a viable method to let AI have a long vision to predict any risks while making any decisions. To make it clear, this paper mainly classifies this scenario into three categories, and all of them will be demonstrated in images. By default, all vehicle symbols in the images are depicted as moving to the right. The orange ones represent vehicles using the RL approach.

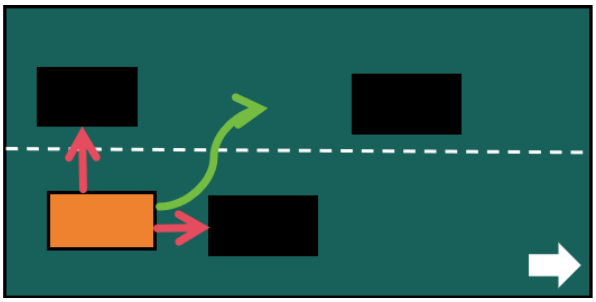


Figure 2. Changing lanes in predictable traffic flow

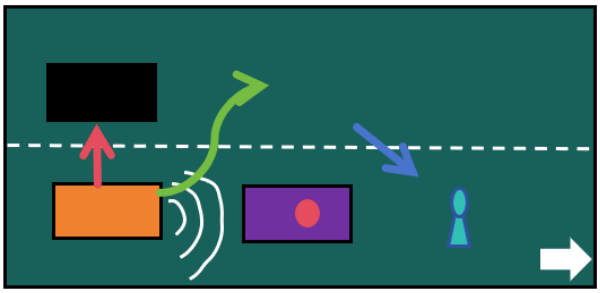


Figure 3. Changing lanes by preventing sudden pedestrian (unpredictable)

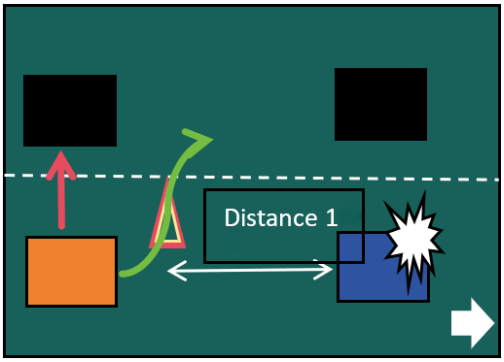


Figure 4. Changing lanes by preventing accidents or obstacles (predictable)

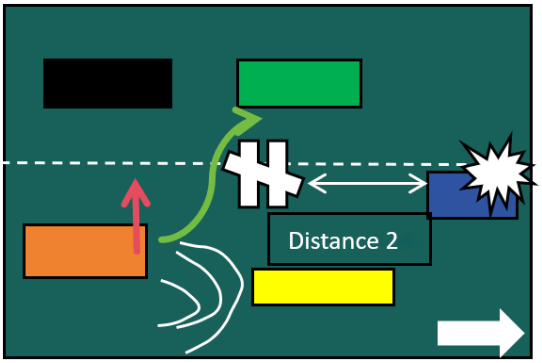


Figure 5. Changing lanes by preventing accidents or obstacles (unpredictable)

In Figure 2, it is the most basic scenario that a car usually faces on a highway. This picture illustrates how this method helps the AI change lanes in a fast-moving traffic flow. There are two arrows marked in red, which represent the AI's active analysis of the distance between the preceding vehicle and another vehicle that is about to enter the lane (this step will exist in every scenario). By applying the fixed processing formula mentioned earlier, it will obtain an accurate piece of information to make a decision and enter the other lane.

$$\text{Decision} = \text{Road condition}*N+C \tag{4}$$

In Figure 3, this can be regarded as the most hazardous driving scenario. Suddenly, the car in front (purple) encounters a pedestrian who appeared out of nowhere. How can the orange car use the RL approach to avoid a potential accident? Firstly, the in-vehicle AI will actively track the driving status of the preceding vehicle, like radar, and maintain an absolutely safe distance to facilitate emergency avoidance. At the same time, it will continuously observe whether there is a safe gap in the adjacent lane to facilitate an emergency lane change. If the preceding vehicle makes an abnormal emergency brake (marked as red circular) instead of following the normal braking procedure, then the AI will autonomously determine the timing to perform an emergency brake and, if necessary (if the braking distance is insufficient), merge into another lane. After entering the other lane, continuously observe the movement route of the pedestrians and appropriately reduce the speed to minimize the impact caused by the blind spots in vision. Finally, avoid the pedestrians and continue driving. During the whole process, AI actively learn to collect driving information (average speed) from other cars. All this data will finally be stored in its CPU. By continuously deepening and strengthening learning abilities, AI can improve its decision-making ability while driving by itself.

In Figures 4 and 5, these two are actually the same scenarios, just in two different forms. One is predictable and another is unpredictable. In Figure 4, there was a certain distance between the preceding vehicle (blue) that had an accident and blocked a lane. However, the driver of this vehicle had already placed a warning sign at an appropriate safe distance (distance 1) in advance. Then, the AI will actively recognize this warning sign and start to appropriately slow down in advance. It will analyze the driving conditions in other lanes and look for an opportunity to make a lane change. But if there are no warning signs (figure 5), then AI will continue to track the driving information of the preceding vehicle as it did in the previous scenario. If it is detected that the speed of the preceding vehicle is slower than that of other lanes and does not conform to normal driving behavior (as shown in the diagram, the speed of the green vehicle is not equal to that of the yellow vehicle), then the possibility of an accident occurring ahead increases. In conjunction with satellite navigation, it will calculate the distance to the potential accident (distance 2) within the processor to make an early lane change.

In short, it only lists the most common scenarios of high-speed lane changes. For specific countries and specific policies, of course, there are many complex situations that this study has not taken into account. However, this method has made the greatest contribution to making the decision-making of autonomous driving systems more reliable and stable in simple risk-avoidance situations.

4. Alleyways

As the last driving scenario listed in this paper, the alleyways can be said to be the most complex of the three. It is different from the first two scenarios and cannot be dealt with comprehensively. Other potential factors (e.g., kids, motorcycles) can also make it very tough for AI autonomous driving systems to operate in such extremely variable scenarios. A study compared the alleys in Beijing, China, with those in Cincinnati, USA (Marinic et al., 2021). In that research, whether from a historical perspective or a geographical one, due to the high population density, it is obvious that the alleys in Beijing are much more complicated and diverse than those in Cincinnati. In historical perspective, although China’s social and economic transformation allowed for significant architectural interventions to reshape the urban structure (Maroldi et al., 2019). Due to the protection

of cultural heritage, these alleys (Hutongs) were preserved and remained intact. Gradually, they became part of urban planning. From a geographic perspective, the overly narrow roads and unclear traffic management have significantly increased the difficulty for vehicles to navigate through these alleys. In this paper, if it merely focuses on summarizing and analyzing the alleys in various nations and religions to optimize the AI in the autonomous driving system, it would be too broad and complicated. RL approach emphasizes that the vehicle AI ought to learn the external environment during the driving process and make correct decisions. Therefore, the RL approach in this part will be optimized and applied from the perspective of the vehicle's subjective driving function.

To understand how it works in detail, this section of the paper will explain how the RL approach enhances the random adaptability and speed control ability of vehicles in urban streets and alleys in one specific country (China) by upgrading the navigation and braking system.

When drivers manually drive their cars through narrow alleys, on-board radars are usually the primary technical means they rely on for early warning and avoidance of obstacles. On the contrary, due to the fact that radars often malfunction in conditions such as insufficient lighting or blocked vision, drivers are unable to accurately judge the external environment, thereby leading to accidents (Valero et al., 2025). Besides, pedestrians can also be a remarkable problem. Olasupo (2022) argues that it is even harder for a human driver to know what decision a pedestrian at a traffic intersection will make, either to enter the road when the pedestrian light is red and the car lights are green in the traffic intersection or not. For AI, it will be a challenge to predict and avoid these unpredictable factors during driving. Therefore, it is essential to install and optimize the RL approach in the navigation module of the autonomous driving system. Another very vital function is the automatic emergency braking (AEB) system. It is considered to be an effective active safety system for avoiding rear-end and pedestrian collisions. This system is an advanced assistance system designed to identify imminent collisions and react by automatically activating the brakes, and is based on camera recognition of an object in front of the vehicle (Tan et al., 2020). In the United States, the Highway Loss Data Institute (HLDI) compared rates of insurance claims per insured vehicle year for Volvo S60 and XC60 vehicles with standard low-speed AEB with comparable vehicles without AEB. Volvos with low-speed AEB had 18% fewer collision claims, which covered damage to the at-fault driver's vehicle. Based on EU crash data, Wilmink (2008) reported that the AEB system can reduce injuries and fatalities by 7% each. It is obvious that the AEB system plays a crucial role in preventing and minimizing the risk of accidents. This method will integrate the vehicle's navigation system and AEB system. Optimizing and modifying them in specific scenarios.

A simple diagram to demonstrate this approach. Taking the alleyways in Beijing, China, as an example.

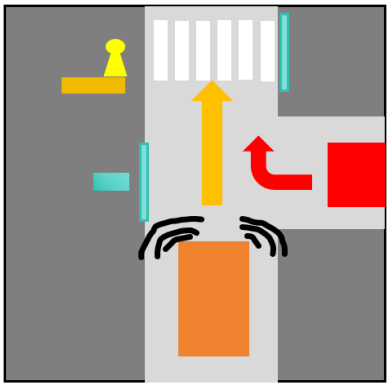


Figure 6. How the RL approach works in avoiding potential risk factors

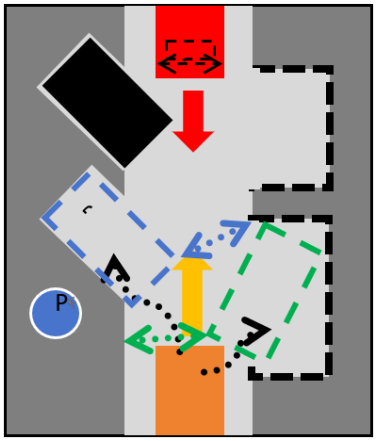


Figure 7. How the RL approach works in facing an oncoming car

These two diagrams each demonstrate a common driving scenario in alleys in Beijing, China. Both of them highlight the importance of the cooperation of the intelligent navigation system and AEB in the RL approach.

In Figure 6, assuming the orange is the sample car (using autonomous driving) and the red is another vehicle. Other features are marked in the diagram as well. Orange's target is to directly go through the T-junction (very popular in alleys in Beijing). Red's target is to turn right at that corner. There are no traffic lights in alleys, and with a large number of random factors (pedestrians, doors may open at any moment). For AI, it needs to plan the driving route in advance and mark every possible intersection on the in-car map. Secondly, before passing each intersection, it uses radar perception to detect whether there are other vehicles (red) that may cause conflicts with its (orange) own route. During this process, it constantly recognizes the roadside buildings and records every house or passage where people may appear. If it encounters an unfamiliar situation, it will record it in its driving memory storage. Finally, the AI will slow down at least to half of the maximum speed limit in the area before each zebra crossing to deal with pedestrians who may rush out from the blind spot. With that low speed and adequate preparation, the success rate of AEB will increase significantly.

In Figure 7, it is considered to be one of the most troublesome situations for drivers across the world. In a narrow alleyway where only one car is allowed to travel in one direction, an opposing car (red) comes from the opposite direction. Luckily, there are some parking spots to use on account of these alleys in Beijing, usually near residential areas. Thus, when the sample car meets this situation, it will first make a judgment whether the road is wide enough for two cars passing in parallel. If not, then it will navigate the routes to avoid the opposing car. As can be seen from the diagram, there are two marked routes, x and y.

Assuming the opposing car width is s (1.7m). When the sample car enters either of the two routes (x and y), the corresponding driving widths reserved for the opposing car are x1 and y1. AI will make its decision by evaluating each of the two distances. It can be shown in a general way.

$$\begin{gathered}\text{If selecting route x, then...} \\ x1\ (1.6m) < s\ (1.7m) \\ \text{Insufficient driving width} \\ \text{If selecting route y, then...} \\ y1\ (1.9m) > s\ (1.7m) \\ \text{Sufficient driving width} \\ \text{Still, this decision will be operated by following the basic formula} \\ \text{Decision} = \text{Road condition}*N+C\end{gathered} \tag{5}$$

Finally, AI will choose route y as it provides enough driving space for the opposing car to drive through.

Admittedly, the structure of alleys in other countries may vary, so the core of this RL approach is to reduce the risks of driving in such complex road conditions through the car's own navigation and speed control system. In another dimension, the alley itself is a very extreme driving environment. Compared with the other two scenarios mentioned above, it is full of random events and obstacles. Motorcycles, garbage bins, or stray cats are all potential factors that may affect the autonomous decision-making ability of AI. Therefore, improving the traffic conditions in the local alleys through relevant departments is also an indispensable part of reducing risks.

5. Conclusion

This paper briefly reviews the current development of autonomous driving systems and presents a method from the specific perspective of AI decision-making. This method emphasizes boosting the autonomous driving performance of AI in driving through different scenarios. This paper generally

provides three common scenarios that drivers usually face every day. Including traffic jams, highway lane changing, and alleyways. AI in each of these scenarios has a different modified version of the RL approach, but following the same basic formula. This reinforcement learning approach puts forward a concept of making AI judge the driving environment and predict any potential dangers. By learning and recording the road conditions driven, it actively enhances AI's own decision-making ability. Compared with current mechanisms of autonomous driving systems on the market that are all confined to certain fixed logic, this approach demonstrates a certain degree of flexibility and universality. It can positively contribute to improving the safety of autonomous driving and alleviating regional traffic pressure, among others.

Nevertheless, the RL approach proposed in this paper also has certain limitations. It does not consider the influence of various road policies in different countries. Precisely, it still lacks a real-life test to reflect whether it can be a crucial and practical function in a car or just a theoretical hypothesis. In any case, the trend of autonomous driving is consistently increasing. If RL approach can work ideally, it will bring optimistic transformation for the intelligent driving system industry and make a certain contribution to driving safety.

References

[1] Bhardwaj, A., Iyer, S. R., Ramesh, S., White, J., & Subramanian, L. (2023). Understanding sudden traffic jams: From emergence to impact. *Development Engineering*, 8, 100105.

[2] Guan, Y., Li, S. E., Duan, J., Wang, W., & Cheng, B. (2018). Markov probabilistic decision making of self-driving cars in highway with random traffic flow: a simulation study. *Journal of Intelligent and Connected Vehicles*, 1 (2), 77 - 84.

[3] Kesting, A., & Treiber, M. (2013). *Traffic flow dynamics: data, models and simulation. No. Book, Whole*) (Springer Berlin Heidelberg, Berlin, Heidelberg.

[4] Mahmud, M., Kaiser, M. S., Hussain, A., & Vassanelli, S. (2018). Applications of deep learning and reinforcement learning to biological data. *IEEE transactions on neural networks and learning systems*, 29 (6), 2063 - 2079.

[5] Marinic, G., Radtke, R., & Luhan, G. (2021). Critical Spatial Practices: A Trans-scalar Study of Chinese Hutongs and American *Alleyways. Interiority*, 4 (1).

[6] Maroldi, F., Lunati, M., Moretti, M., & Pelucchi, G. (2019). Reuse Design in Historical Context: TheHutong Areas in Beijing. *In Edulearn19 Proceedings* (pp. 10226 - 10235). IATED.

[7] Mohammad Hasani, A., Mehrivash, H., Lynch, A., & Shu, Z. (2021). *Reinforcement learning based safe decision making for highway autonomous driving.* arXiv preprint arXiv: 2105.06517.

[8] Olasupo, O. E. (2022). Review of challenges and proposed research directions and radar sensor solution for the deployment of self-driving cars. *International Journal of Advanced Networking and Applications*, 14 (3), 5471 - 5476.

[9] Rahman, R. A., Radzi, A. R., Saad, M. S. H., & Doh, S. I. (2020). Factors affecting the success of highway construction projects: the case of Malaysia. *In IOP Conference Series: Materials Science and Engineering* (Vol. 712, No. 1, p. 012030). IOP Publishing.

[10] Tan, H., Zhao, F., Hao, H., Liu, Z., Amer, A. A., & Babiker, H. (2020). Automatic emergency braking (AEB) system impact on fatality and injury reduction in China. *International journal of environmental research and public health*, 17 (3), 917.

[11] Valero, J. G., González-López, G., Expósito, A., Estrada, L. V., & Lobo, F. J. (2025). *Automotive In-Cabin Radar Uncovered: The Essential Guide to Choose the Perfect Sensing Technology for Your Vehicle*. IEEE Transactions on Microwave Theory and Techniques.

[12] Wilmink, I.; Janssen, W.; Jonkers, E. (2008) *Socio-economic Impact Assessment of Stand-alone and Co-operative Intelligent Vehicle Safety Systems (IVSS) in Europe*. Impact assessment of Intelligent Vehicle Safety Systems Report D4.