Highway decision-making strategy for autonomous vehicle for overtaking maneuver using deep reinforcement learning (DRL) method

Ali Rizehvandi, Shahram Azadi*

Faculty of Mechanical Engineering, K.N.Toosi University of Technology, Tehran, Iran

ABSTRACT

Automated driving represents a novel technology aimed at reducing traffic accidents and enhancing driving efficiency. This research introduces a deep reinforcement learning (DRL) approach for autonomous vehicles, focusing on overtaking scenarios on highways. Initially, a highway traffic environment is established, to guide the agent through surrounding vehicles both efficiently and safely. A hierarchical control framework is outlined to manage high-level driving decisions alongside low-level control aspects like car speed and acceleration. Subsequently, a specialized DRL-based method known as Deep Deterministic Policy Gradient (DDPG) is employed to devise decision-making strategies on the highway. The DDPG offers continuous action space exploration, making it suitable for tasks like autonomous driving where actions are not discrete. Unlike DQN, it can handle high-dimensional action spaces more effectively, enhancing its applicability in complex environments like highway driving. The efficacy of the DDPG algorithm is compared to that of the DQN algorithm, with subsequent evaluation of the results. Simulation outcomes demonstrate that the DDPG algorithm adeptly handles highway driving tasks with efficiency and safety. The study underscores the potential of DRL techniques, particularly the DDPG approach, in advancing the capabilities of autonomous vehicles and improving their performance in complex driving scenarios.

KEYWORDS

Autonomous vehicles, decision making, DRL method, overtaking, DDPG algorithm.

* Corresponding Author: Email: Azadi@kntu.ac.ir

1. Introduction

Autonomous driving enables a vehicle to participate in various driving scenarios without human intervention. Given the vast potential of artificial intelligence (AI), self-driving vehicles have become a major focus of research worldwide. Then, researchers in the automotive sector are aiming to build highly advanced self-driving cars. In recent years, numerous studies have been conducted on autonomous driving based on the deep reinforcement learning (DRL) method. For example, Duan and colleagues proposed a hierarchical structure for learning decision-making policies through the reinforcement learning (RL) method [1]. In recent years, the use of deep reinforcement learning (DRL) algorithms with continuous action spaces for decision-making processes in autonomous vehicles has become widespread. For instance, in [2], an actor-critic algorithm based on reinforcement learning (RL) was used to learn the decision-making process of an autonomous vehicle on a highway. Moreover, in [3], the deep deterministic policy gradient (DDPG) algorithm was studied for continuous decision-making of an autonomous vehicle in an urban intersection environment.

In this research, a driving policy based on a deep reinforcement learning (DRL) approach is proposed for overtaking maneuvers in highway traffic environments for autonomous vehicles. The proposed decision-making strategy ensures safety and efficiency in complex scenarios.

2. Methodology

In this section, we introduce the driving scenario studied on the highway, specifically the important and common overtaking scenario. For designing the scenario, the MATLAB 2022 software environment was used. First, a three-lane highway environment was constructed. Then, the agent and the surrounding traffic environment were designed. Additionally, a hierarchical motion controller was introduced to manage the lateral and longitudinal movements of the agent and the surrounding vehicles.

Decision-making in autonomous driving involves selecting a sequence of logical driving behaviors to achieve specific driving goals. In the highway overtaking maneuver, these behaviors include lane changing, lane keeping, accelerating, and braking. The main objectives are avoiding collisions, moving efficiently and quickly, and driving in the fast lane. In other words, accelerating and overtaking other vehicles is a common driving behavior known as overtaking. This work discusses the decision-making problem on the highway for autonomous vehicles, and the driving scenario is shown in Figure 1. The orange vehicle represents the agent, and the green vehicles are referred to as surrounding vehicles. The agent starts driving in the middle lane at a random speed. The goal of the agent is to drive at the highest possible speed without colliding with the surrounding vehicles. Hence, these goals translate to the efficiency and safety of the decision-making algorithm.

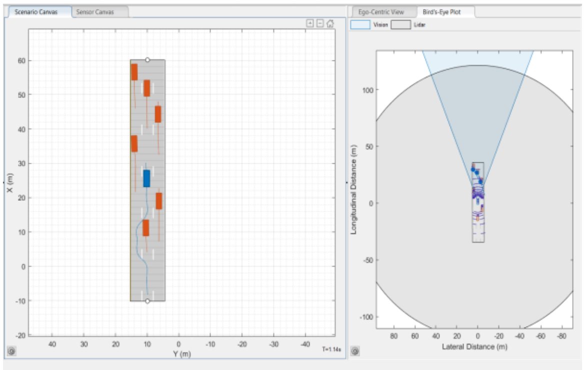

Figure 1. Highway traffic environment design in MATLAB software (2022 version)

It is assumed that the position, speed, and acceleration of the surrounding vehicles are known to the agent. These constraints drive the agent to learn to drive in the designed scenario through a trial-and-error method.

To extract the decision-making strategy based on DDPG, the variables for simulating the driving environment are initialized as follows:

The control actions are the throttle valve and steering angle of the vehicle. Additionally, the state variables (according to equations 1 and 2) are the relative distance and velocity between the agent and the surrounding vehicles:

$$\Delta s = \left| s_{ag} - s_{su} \right| \tag{1}$$

$$\Delta V = \left| V_{ag} - V_{su} \right| \tag{2}$$

Here, S and V represent the position and speed information obtained from the vehicle dynamics. Also, the indices ag and su denote the agent and surrounding vehicles, respectively. It's worth mentioning that equations (1) and (2) can also be considered as the transition model PP in the reinforcement learning (RL) framework. Finally, the reward function R in this study consists of three components representing efficiency, safety, and driving objectives. Essentially, the agent must drive at the maximum speed possible, stay in the lane, and avoid collisions with other surrounding vehicles. The designed reward in each time step (t) is defined as follows:

$$R_t = -100(collision) - 40(L-1)^2 - 10(V_{ag} - V_{ag}^{\max})^2 \quad (3)$$

Here, collision is defined as {0,1}, indicating collision conditions for the agent. Also, the number of lanes is denoted by {1,2,3}, representing the lane number on the highway. In this study, the proposed decision control policy is simulated, trained, and evaluated in the MATLAB environment. The number of lanes and surrounding vehicles are set to 3 and 6, respectively. Additionally, the discount factor and learning rate are 0.8 and 0.2, respectively. The value of the value network layers is set to 128. Also, the number of episodes is set to 100. Figure 2 shows the block diagram of the simulated overtaking scenario in the MATLAB-Simulink software.

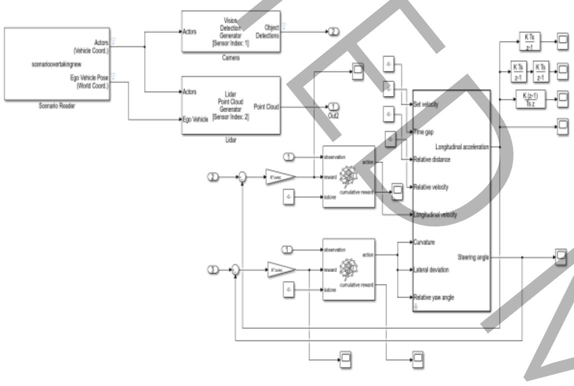

Figure 2. Block diagram of overtaking scenario in Simulink software

According to Figure 2, the block diagram consists of three subsystems: the scenario, reinforcement learning, and vehicle dynamics along with its control. In the next section, the performance of the decision-making algorithm presented in the reinforcement learning subsystem will be discussed.

3. Results and Discussion

In this section, the control performance of the proposed DDPG algorithm for the decision-making process of the agent in the highway traffic environment is evaluated and analyzed. In this section, two methods for agent decision-making in the highway traffic environment are presented. Firstly, the deep Q-learning algorithm (DQN) is examined, followed by the proposed DDPG algorithm, which is evaluated and analyzed. Additionally, the advantages and superiority of the DDPG algorithm over the deep Q-learning algorithm (DQN) are demonstrated in this section. Essentially, the deep Q-learning algorithm (DQN) serves as a benchmark for inferring the optimality of the DDPG algorithm. The average rewards for both deep Q-learning (DQN) and DDPG methods over 25 episodes are shown in Figure 3.

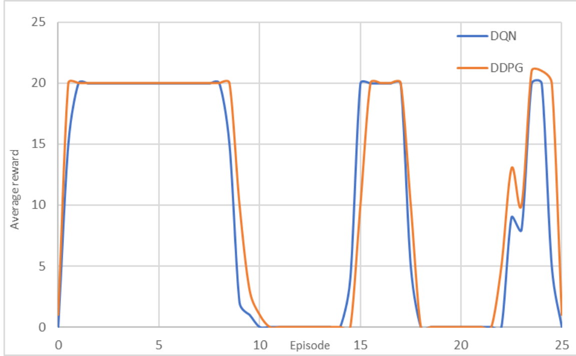

Figure 3. Average reward in DDPG and DQN methods

The increasing trend of these curves indicates better performance of the agent in interacting with the environment. Additionally, the decreasing trend of the curves is due to the probability of the agent colliding with surrounding vehicles and crossing existing lanes on the highway during the overtaking maneuver. This is a result of the complex traffic environment designed on the highway. According to Figure 3, the learning rate of the DDPG algorithm for performing the designed overtaking maneuver in the highway environment is better than the DQN algorithm.

4. Conclusion

In this research, an efficient and safe decision-making algorithm based on deep reinforcement learning (DRL) method on the highway for autonomous vehicles is proposed. The mentioned algorithm is DDPG. Simulation results demonstrate that the proposed decision-making algorithm can ensure optimality and convergence rate. Additionally, the proposed algorithm's learning rate is higher than the DQN algorithm.

5. References

[1] Duan, J., Li, S. E., Guan, Y., Sun, Q., & Cheng, B., Hierarchical reinforcement learning for self-driving decision-making without reliance on labeled driving data, IET Intelligent Transportation Systems, 14(5), (2020) 297-305.

[2] Duan, J., Guan, Y., Li, S. E., Ren, Y., & Cheng, B., Distributional Soft Actor-Critic: Off-Policy Reinforcement Learning for Addressing Value Estimation Errors, IEEE Transactions on Neural Networks and Learning Systems, 33(5), (2022) 2345-2357.

[3] Li, G., Li, S., Li, S., & Qu, X., Continuous decision-making for autonomous driving at intersections using deep deterministic policy gradient, IET Intelligent Transportation Systems, 16(2), (2021) 1669-1681.