Interactive Driving Control Design for Autonomous Vehicles Using Deep Reinforcement Learning

Mei-Lin Huang and Ching-Hung Lee
Institute of Electrical and Control Engineering
National Yang Ming Chiao Tung University
Hsinchu City, Taiwan
adnil.huangn2@gmail.com; chl@nycu.edu.tw

Hsin-Han Chiang
Department of Vehicle Engineering
National Taipei University of Technology
Taipei, Taiwan
hsinhan@mail.ntut.edu.tw

***Abstract*—The decision making and motion planning continue to pose many challenges to the reliability and safety of autonomous vehicles (AV), especially under complex and nondeterministic traffics. Nowadays, growing research on AV attempts to solve this problem by means of reinforcement learning paradigms while endowing machines with how to behave through environmental interaction. In this study, a deep reinforcement learning approach with a modified version of the deterministic policy gradient (DDPG) is proposed to implement an interactive driving control for AV. Under the developed co-simulation framework, the deep controller to achieve interactive driving safety of AV can be generated automatically with the aid of model-free deep reinforcement learning.**

***Keywords*—Interactive driving, reinforcement learning, deep neural networks, path tracking, autonomous vehicles.**

I. Introduction

Autonomous vehicles (AV) have been an emerging application worldwide in recent years due to the expected improvement in traffic safety and fuel consumption. AV should possess several abilities to perform interactive driving. The first ability is to recognize the driving environments by using different types of sensors such as lidars, radars, and cameras. The second ability is to keep collision avoidance in the longitudinal direction well, and the third ability requires guiding AV along reference trajectories at all admissible speeds even in the presence of disturbances. Thus, designing a safe and reliable decision and planning system for AV in diverse traffic environments is particularly challenging involving many difficulties from several aspects [1], such as the complex conditions of roads, including geometry, topology, and road marks in different scenarios. Moreover, to handle the unknown future motion of multi-agents on the road, the autonomous driving system also needs the requirement of fast computation to react to highly dynamic surrounding objects and potential emergencies.

Due to the large complexity and difficulty of creating model-based autonomous driving controllers for AV, recent researchers in academia and industry have attempted to solve this critical problem with reinforcement learning methods [2, 3]. One of the most commonly used input types of deep learning methods is an image. However, images are data with high dimension, though it includes information we need, it also includes information that is not necessary and often affects the performance of the model. For instance, brightness, sky, buildings, and so on. Not only does it affect the performance, but also the complexity makes the model difficult to converge. Besides, deep learning of image recognition is easy to fool [4]. To address these issues, this study employs the deep deterministic policy gradient (DDPG) algorithm to make autonomous cars learn reliable and safe driving behaviors. To perform safe interactive driving with multiple vehicles in the simulated traffic flow, our proposed approach includes a new input representation and adopts deep reinforcement learning into the design of a model-free deep controller.

This work was financially supported by National Science and Technology Council, Taiwan, under Grants No. 111-2221-E-027-147-MY2.

II. Method

The deep controller architecture proposed in this study includes two planners: a speed planner and a steering planner. Both planners are designed based on the deep learning network. This section introduces the architecture with the input representation and reward function design.

A. Input Representation

The inputs to the network include the reference path, information on surrounding vehicles, and the current speed of the ego vehicle. By getting the trajectory points in terms of reference path during a sampled interval, these points will go through a feature extraction model and then output the four parameters, as shown in Fig. 1. The information of surrounding vehicles includes relative velocity and relative position in the x-axis and y-axis separately. Each axis of velocity and position will be formed into a 3 by 3 matrix. Each cell in a matrix represents the direction of surrounding objects, including forward, right-forward, right, right-rear, rear, left-rear, left, and left-forward.

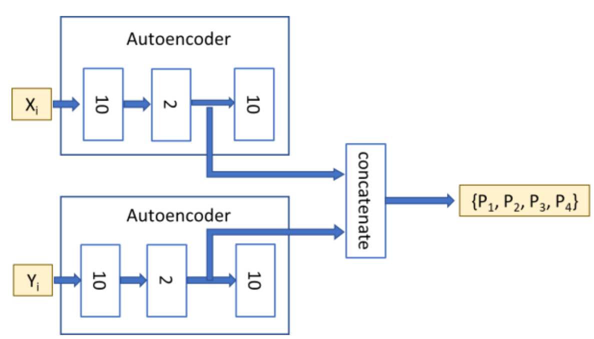

Fig. 1. Path feature extractor for the driving controller.

B. Control Design

DDPG is a reinforcement learning method that adapts the ideas underlying the success of deep Q-learning to the continuous action domain [5]. With an actor-critic scheme, the model-free algorithm in DDPG based on the deterministic

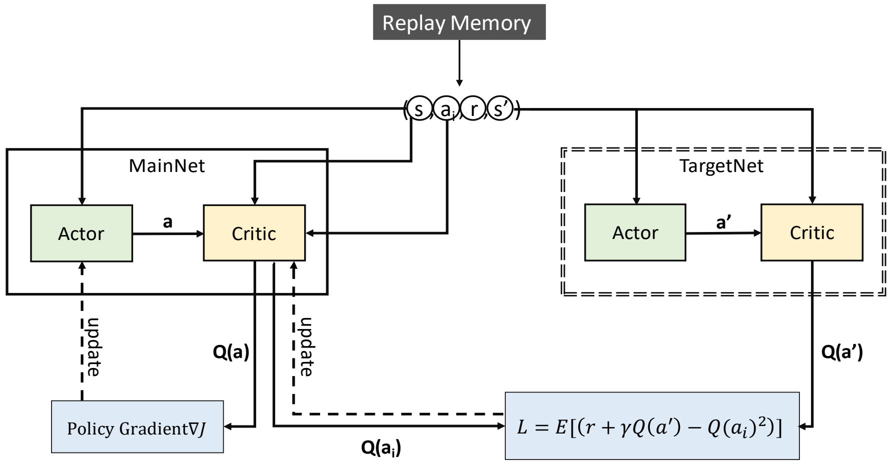

Fig. 2. The proposed deep learning-based controller architecture.

policy gradient can operate over continuous action spaces. Fig. 2 shows the designed architecture of the deep controller, in which s represents the state including the current speed of the ego vehicle, information of surrounding objects, relative speed and relative position, and reference path; a is the action that represents the target speed and target steering angle; r represents the reward received from the environment.

C. *Reward Function*

There are two planners in the deep controller, and the associated two reward functions are given as:

$$\mathrm{r}_{tv} = \mathrm{r}_b + \mathrm{r}_v + \mathrm{r}_a + \mathrm{r}_d \tag{1}$$

$$\mathrm{r}_{sw} = \mathrm{r}_y + \mathrm{r}_l \tag{2}$$

where r_{tv} represents a reward for the target speed planner and r_{sw} represents a reward for the target steering angle planner. r_b is the term penalizing collision with surrounding objects, if a collision occurs, then $\mathrm{r}_b = -10$, otherwise $\mathrm{r}_b = 0$. r_v is the term encouraging moving forward and r_v is set to the current speed v if $0 < v \leq 50$. Moreover, to avoid over-speeding or staying still $\mathrm{r}_v = -5$ if $v > 50$ and $\mathrm{r}_v = -1$ if $v = 0$. r_a is the term penalizing acceleration a, and we set $\mathrm{r}_a = -|a|$, which makes the vehicle perform stable driving in traffic. r_d is the term penalizing distance d with the preceding car, if d is farther than the specified safety distance then $\mathrm{r}_d = 0$, otherwise $\mathrm{r}_d = -1$. r_l is the term penalizing going out of the lane, where $\mathrm{r}_l = -2$ if the lateral offset is more than 1.75m, otherwise $\mathrm{r}_l = 0$. To improve the motion stability of the ego vehicle, $\mathrm{r}_y = -|\omega|$ where ω is the yaw rate of the ego vehicle.

III. Data Collection and Simulation Environment

This study trains and evaluates the proposed method by using the co-simulation between the CarSim simulator and Python software. The CarSim simulator can support accurate, detailed, and efficient methods for simulating the performance of vehicle driving, and also provides sensing data such as relative distance between vehicles and lane information of the subject vehicle. The co-simulation schematic built in this study is given in Fig. 3. During the training stage, several traffic vehicles are randomly placed around the ego vehicle, and their speeds are also varied randomly so that diverse and sufficient traffic scenarios can be considered in training and evaluation. Besides, some traffic vehicles can perform lane-change or cut-in behavior, making the traffic quite challenging. The proposed deep controller will be evaluated under varying traffic densities in order to demonstrate the capability of our approach in the safe interactive driving task.

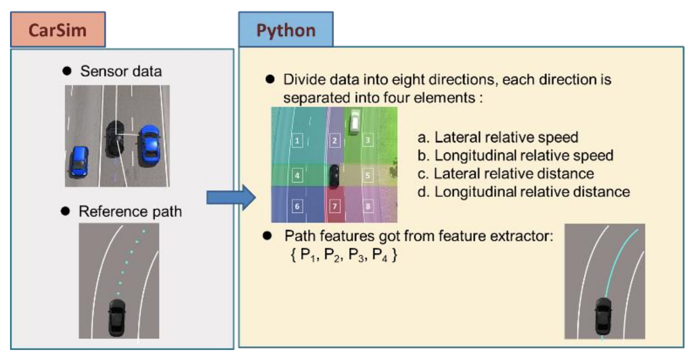

Fig. 3. Co-simulation framework used for training and testing the proposed deep controller in the experimental scenario.

References

[1] L. G. Cuenca, E. Puertas, J. F. Andrés, and N. Aliane, "Autonomous driving in roundabout maneuvers using reinforcement learning with Q-learning," *Electronics*, 8(12), 1536, 2019.

[2] Chen J., Yuan B., and Masayoshi T., "Deep imitation learning for autonomous driving in generic urban scenarios with enhanced safety," arXiv:1903. 00640v1, 2019.

[3] S. Nageshrao, E. Tseng, and D. Filev, "Autonomous highway driving using deep reinforcement learning," arXiv:1904.00035v1, 2019.

[4] Douglas H., "Deep trouble for deep learning", *Nature*, vol. 574, pp. 163-166, 2019.

[5] Timothy P. L., Jonathan J. H., Alexander P., Nicolas H., Tom E., Yuval T., David S. Daan W., "Continuous control with deep reinforcement learning", arXiv:1509.02971, 2015.