Pre-training Framework for Improving Learning Speed of Reinforcement Learning based Autonomous Vehicles

Jung-Jae Kim
POSCO R&D Center
Pohang, Korea
jungjae.kim@posco.com

Si-Ho Cha*
Dept. of Multimedia Science
Chungwoon University, Incheon, Korea
shcha@chungwoon.ac.kr

Minwoo Ryu
Service Laboratory of Convergence Technology
KT R&D Center, Seoul, Korea
my.ryu@kt.com

Minho Jo
Dept. of Computer Convergence Software
Korea University, Sejong, Korea
minhojo@korea.ac.kr

***Abstract*—Reinforcement learning based autonomous vehicles have the disadvantage of long learning time. The paper proposes a pre-training framework for improving learning speed of autonomous vehicles (PRELSA) in reinforcement learning. PRELSA framework pre-learns the agent's neural network before actual learning by pre-initializing the agent's policy gradient neural network. Simulation results show that PRELSA framework improves learning speed by about 20 percent compared to existing learning method.**

Keywords*—*autonomous vehicles; reinforcement learning; pre-training framework; learning speed

I. Introduction

Autonomous vehicles can recognize the surrounding environment and analyze the data exchanged between them. They consist of four modules: sensing, recognition, decision making, and control [1].

An autonomous vehicle recognizes the data sensed by sensors as an entity and makes autonomous driving by controlling the parts of the vehicle by making a decision based on the data [2]. It is important for the decision making module to determine the current situation and make a decision. However, existing autonomous vehicles can't cope with new situations and cause policy conflicts. Therefore, decision making modules based on reinforcement learning have been studying.

However, reinforcement learning can only determine the value of the space after the agent has explored the state space, which means that in reinforcement learning, the agent must search all space for optimal behavior [3]. Therefore, when reinforcement learning is applied to autonomous vehicles, it takes longer learning time (search time) than rule based or policy based schemes.

In this paper, we propose a pre-training framework for improving learning speed of autonomous vehicles (PRELSA) in order to overcome the shortcomings of long learning time of reinforcement learning. The proposed RRELSA framework pre-learns the neural network before actual learning by pre-initializing the agent's policy gradient neural network based on the extracted vehicle behavior from pre-learned images or images processed by human experts

* Corresponding author: Si-Ho Cha

II. PRELSA Framework

PRELSA is a framework to initialize the agent's neural network through pre-learning for rapid environment adaptation of autonomous vehicle agents. PRELSA consists of two steps: action extraction and neural network learning.

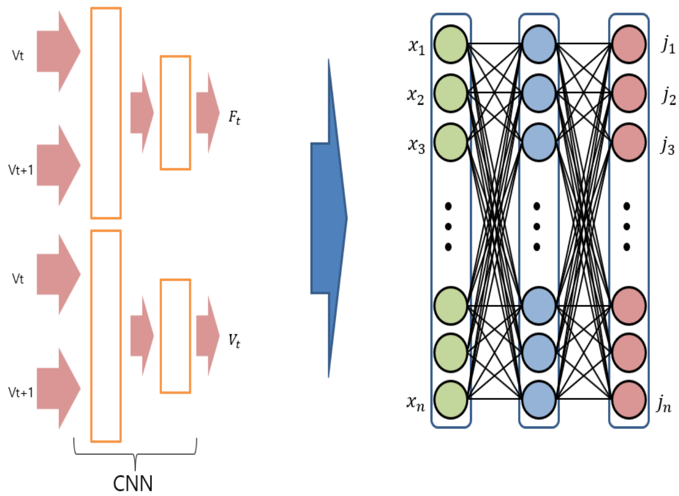

Fig. 1. Basic Process of PRELSA

First, the behavior extraction is done using a convolutional neural network (CNN). The CNN is a neural network that is often used to detect objects in deep neural networks. As shown in Figure 1, PRELSA uses two CNNS and their input are the current frame image and the next frame image. Two CNNS output the speed and direction for the vehicle agent, respectively. In this way, it is possible to extract a series of behavior information of the vehicle agent from the images.

Second, it is neural network learning inside the agent. Since the learning of the neural network inside the agent is the same as that of the feedforward neural network, this paper defines only the input and output to the neural network. The input of the neural network is the environment surrounding the object that can be seen in the image. And the target that the neural network should output is a series of behavior information extracted from CNN. It is the same as a pair of state and behavior required for learning in reinforcement learning.

III. Experiment

In order to evaluate the performance of PRELSA, autonomous driving environment was constructed as an asset of Unity [4]. As shown in Figure 2, the roads consist of bottleneck roads where autonomous driving agents can make the right decisions to make their way out of the road.

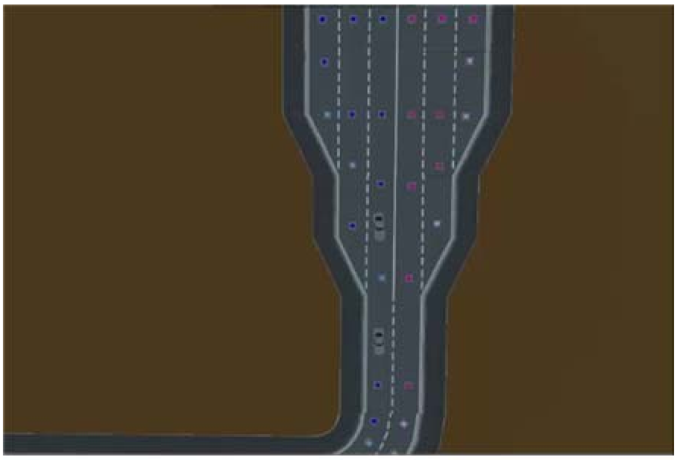

Fig. 2. Bottleneck for evaluation of PRELSA

The roads consisted of a bottleneck structure in which lanes were reduced from 6 lanes to 2 lanes, and only one road was used for the simulation. The simulation process learns that the autonomous vehicle agent can exit the bottleneck road without escaping the road or center line. The minimum speed to exit the road without making a lane change in the center lane is 14 seconds. Therefore, the autonomous driving agent receives a negative reward (14 seconds – current arrival time). Also, vehicle agents are randomly placed in 1 to 3 lanes for each learning. Table 1 compares the convergence time between the existing reinforcement learning method, Actor Critic [5], and PRELSA. A total of 15,000 episodes were trained.

TABLE I. Performance Evaluation

	Convergence Time	*Minimum Escape Time*
Actor Critic	13,120 episodes	10,364 episodes
PRELSA	10,902 episodes	9,098 episodes

As shown in Table 1, the existing reinforcement learning converges after 13,120 episodes, however, when PRELSA framework, only 9,098 training episodes reached the minimum transit time, and after that, 10,902 training episodes showed a stable transit time. This shows that the learning speed using PRELSA framework is improved by about 20%.

IV. Conclusion

In this paper, we proposed a PRELSA framework to pre-learn policy neural networks based on pre-learned images or images of human experts in order to solve problems related to slow learning speed when applying reinforcement learning to autonomous vehicles. Simulation results show that PRELSA framework improves learning speed by about 20% compared to existing learning method, Actor Critic. However, PRELSA framework only extracts the direction and speed of the vehicle agent, and has the disadvantage that the input vector must be input manually. Therefore, it is necessary to study to extract the input vector from the images.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1A09917662).

References

[1] Bob Touchton, Tom Galluzzo, Danny Kent, Carl Crane, “Perception and planning architecture for autonomous ground vehicles”, IEEE Computer, 2006.

[2] Jess Levinson, Jake Askeland, Jan Becker, Jennifer Dolson, David Held, Soeren Kammel, J. Zico Kolter, Dirk Langer, Oliver Pink, Vaughan Pratt, Michael Sokolsky, Ganymed Stanek, David Stavens, Alex Teichman, Moritz Werling, Sebastian Thrun, “Towards fully autonomous driving: Systems and algorithms”, IEEE Intelligent Vehicles Symposium (IV), 2011.

[3] Chenyi Chen, Ari Seff, Alain Kornhauser, Jianxiong Xiao, “DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving”, The IEEE International Conference on Computer Vision (ICCV), pp. 2722-2730, 2015.

[4] Unity, https://unity3d.com/kr/machine-learning.

[5] Vijay R. Konda, John N. Tsitsiklis, “Actor-Critic Algorithms”, Advances in neural information processing systems, pp. 1008-1014, 2000.