Human-Like Autonomous Car-Following Planning by Deep Reinforcement Learning

Meixin Zhu
Graduate Research Assistant
School of Transportation Engineering
Tongji University
4800 Cao'an Road, Jiading District
Shanghai, 201804, China
Email: 1151208@tongji.edu.cn

Xuesong Wang, Ph.D.
Professor
School of Transportation Engineering
Tongji University
4800 Cao'an Road, Jiading District
Shanghai, 201804, China
National Engineering Laboratory for Integrated Optimization of Road Traffic and Safety Analysis Technologies
88 Qianrong Rd, Wuxi，Jiangsu，214151, China
Phone: +86-21-69583946
Email: wangxs@tongji.edu.cn
(Corresponding author)

Yinhai Wang, Ph.D.
Professor
Department of Civil and Environmental Engineering
University of Washington
Seattle, WA 98195-2700
Email: yinhai@uw.edu

November 2017
Presentation at the TRB Annual Meeting

1 INTRODUCTION

Autonomous driving technology is capable of providing driver convenience and enhancing safety by avoiding some crashes caused by driver errors (*1*). While building autonomous driving algorithms, one significant challenge is to make the autonomous vehicles able to emulate human drivers' intelligence and driving styles while staying within the safety bounds, i.e., human-like driving (*2*-*4*).

We propose a deep reinforcement learning (deep RL) based car-following model that applies neural network and reinforcement learning to address the planning of human-like autonomous car following. Deep RL is a field that seeks to combine the advances in deep neural networks with reinforcement learning algorithms to create agents capable of acting intelligently in complex environments (*5*), and exciting breakthroughs have been witnessed, like deep *Q*-network (*6*) and AlphaGo (*7*).

Figure 1 shows the schematic diagram of human-like car-following framework based on deep RL. During the manual driving phase, data are collected and stored as historical driving data in the database. These data are then fed into a simulation environment, where an RL agent learns from trying and interaction, with a reward function signaling how much the agent deviates from the empirical data. An optimal policy (car-following model) that maps from speed, relative speed, and spacing to following-vehicle acceleration in a human-like way is finally obtained, and can be continuously updated when more data are fed in. This optimal policy will act as the executing policy in the autonomous driving phase.

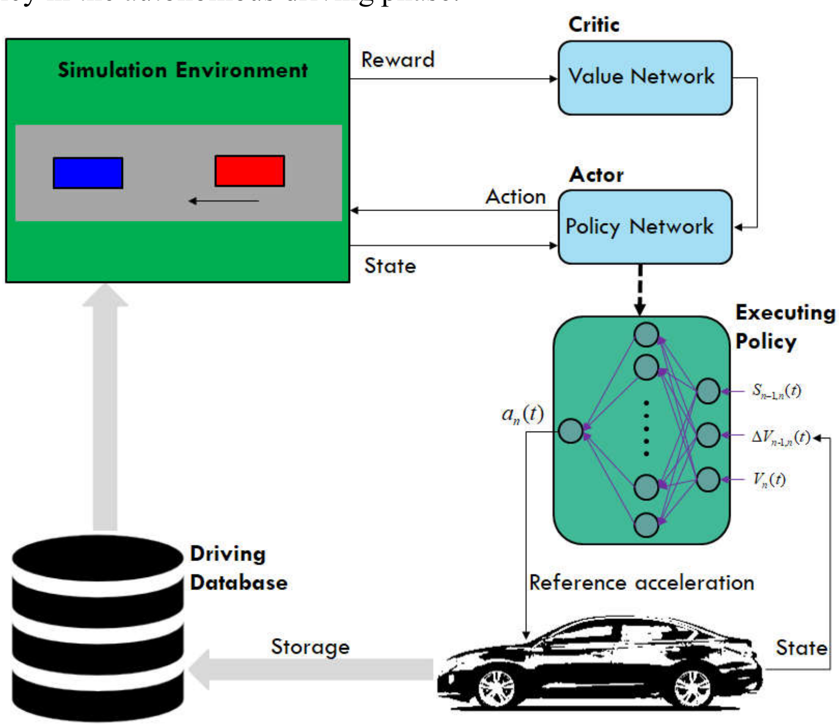


FIGURE 1 Conceptual diagram of human-like car following by deep RL.

To evaluate the proposed model and compare its performance with that of traditional car-following models, real-world driving data collected in the Shanghai Naturalistic Driving Study (*8*) were used to train and test the proposed deep RL model and also several typical traditional car-following models. Their performance in terms of trajectory-reproducing accuracy, generalization, and adaptivity was then compared.

2 METHODOLOGY

Because the acceleration of a vehicle is continuous, a deep RL method called deep deterministic policy gradient (DDPG) (*9*), which performs well in continuous action space, was used. Two separate neural networks were used to represent the actor and critic, respectively. At time step *t*, the actor network takes a state $s_t = (v_n(t), \Delta v_{n-1,n}(t), \Delta S_{n-1,n}(t))$ as input and outputs a continuous action: the following vehicle acceleration $a_n(t)$. The critic network takes as input a state s_t and an action a_t and outputs a scalar *Q*-value $Q(s_t, a_t)$.

Both the actor and critic networks have three layers: an input layer taking the input signals to the whole neural network, an output layer generating the output signal, and a hidden layer containing 30 neurons between the former two layers. We had tested even deeper neural networks, but experiment results showed that considering neural networks deeper than one hidden layers was unnecessary for our problem, which has only three or four input variables.

The hyperparameters (parameters whose values are set prior to the beginning of the learning process) adopted are shown in Table 1. The values of these hyperparameters were selected according to Lillicrap et al. (*9*) and also by performing a test on a randomly sampled training dataset.

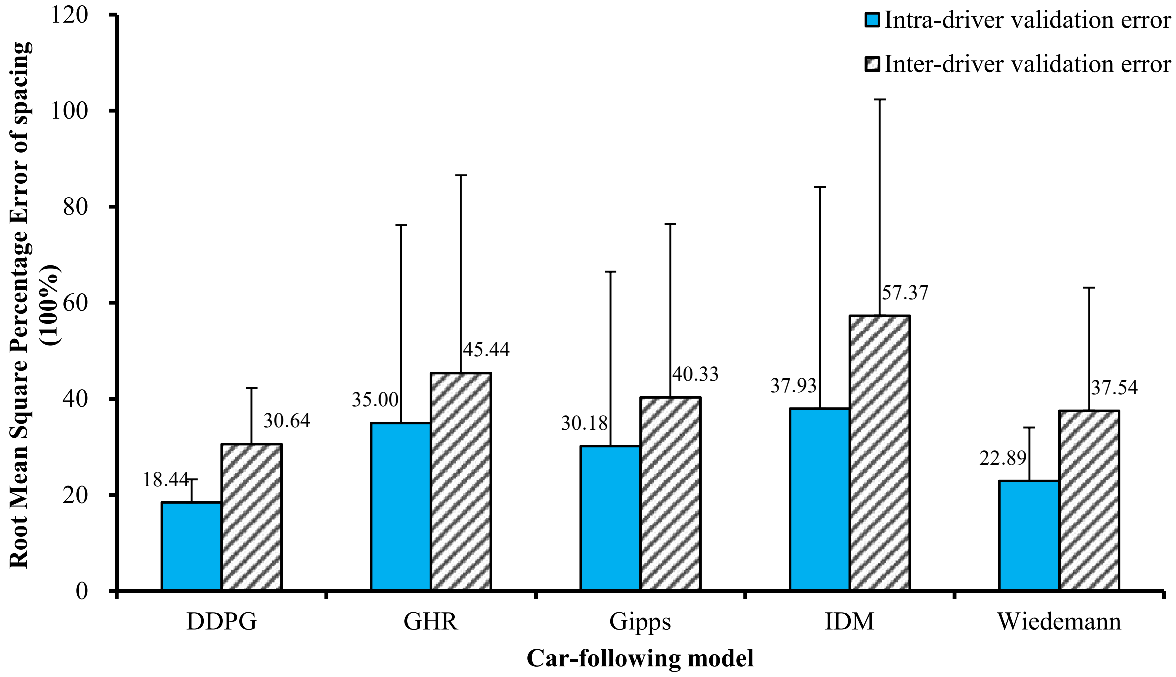
TABLE 1 Hyperparameters and Corresponding Descriptions

Hyperparameter	Value	Description
Learning rate	0.001	Learning rate used by Adam
Discount factor	0.99	Discount factor gamma used in the Q-learning update
Minibatch size	32	Number of training cases over which each stochastic gradient descent update is computed
Replay memory size	7000	Number of training cases in replay memory
Replay start size	7000	Random policy run for this number of time steps before learning starts; the resulting experience is used to populate the replay memory
Soft target update τ	0.001	Update rate of target networks

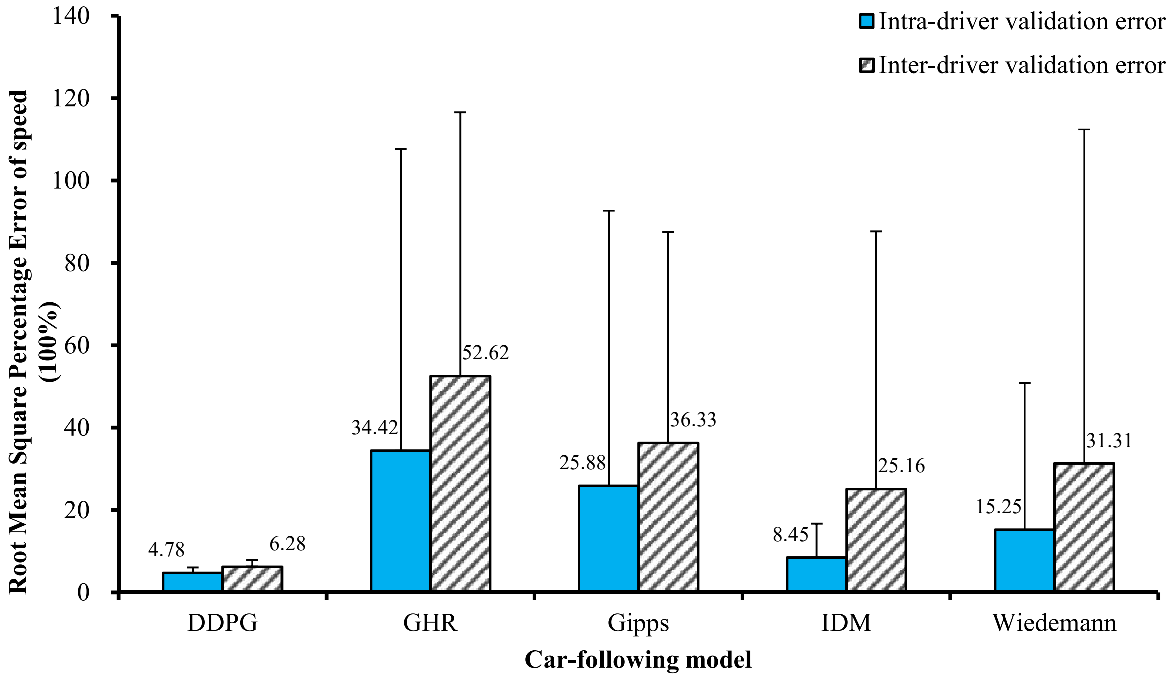
3 RESULTS

Results show that this new model can reproduce human-like car-following behavior with significantly higher accuracy than traditional car-following models, especially in terms of speed replicating. Specifically, the model has a validation error of 18% on spacing and 5% on speed, which is generally 15 and 30 percentage points less, respectively, than that of traditional car-following models. Moreover, the model demonstrates good capability of generalization to different driving situations and can adapt to different drivers by continuously learning.

Figure 2 presents the mean values and standard deviations of the intra-driver validation errors for the five models. The DDPG model outperformed all of the investigated traditional car-following models, as evidenced by its having the lowest mean and standard deviation of errors on both spacing and, especially, speed.



(a) Spacing



(b) Speed

FIGURE 2 Mean standard deviation (line over the bar) of intra- and inter-driver validation errors on (a) spacing and (b) speed for the five models.

4 CONCLUSION

This study has shown the feasibility of a deep RL based car-following model for autonomous driving planning. Real world driving data from SH-NDS was used to train the model and compare its performance with that of traditional car-following models. The results can contribute to the development of human-like autonomous driving algorithms. Moreover, this study demonstrates that data-driven modeling and reinforcement learning methodology can contribute to the development of traffic flow models and offer deeper insight into driver behavior.

ACKNOWLEDGEMENTS

The authors are grateful to Marilyn Morgan and Barbara Rau Kyle for their helpful edit. This study was jointly sponsored by the Chinese National Science Foundation (51522810), the 111 Project (B17032), and the National Engineering Laboratory for Integrated Optimization of Road Traffic and Safety Analysis Technologies.

REFERENCES

1. Wei, J., J. M. Dolan, and B. A. Litkouhi. Learning-Based Autonomous Driver: Emulate Human Driver's Intelligence in Low-Speed Car Following. *Proceeding. of SPIE*, Vol. 7693, pp. 76930L–1.

2. Gu, T., and J. M. Dolan. Toward Human-Like Motion Planning in Urban Environments. *Intelligent Vehicles Symposium Proceedings*, 2014, pp. 350–355.
3. Lefèvre, S., A. Carvalho, and F. Borrelli. Autonomous Car Following: A Learning-Based Approach. *Intelligent Vehicles Symposium*, 2015, pp. 920–926.
4. Kuderer, M., S., and W. Burgard. Learning Driving Styles for Autonomous Vehicles From Demonstration. Robotics and Automation, *IEEE International Conference*, 2015, pp. 2641–2646.
5. Hausknecht, M., and P., Stone. Deep Reinforcement Learning in Parameterized Action Space. International Conference on Learning Representations, 2016.
6. Mnih, V., K., Kavukcuoglu, D., Silver, A. A., Rusu, J., Veness, M. G., Bellemare, and S., Petersen. Human-Level Control through Deep Reinforcement Learning. *Nature*, Vol. 518, No. 7540, 2015, pp. 529–533.
7. Silver, D., A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, and Dieleman, S. Mastering the Game of Go With Deep Neural Networks And Tree Search. *Nature*, Vol. 529 No.7587, 2016, pp. 484–489.
8. Zhu, M., X. S. Wang, A, P, Tarko. Calibrating Car-Following Models on Urban Expressways for Chinese Drivers Using Naturalistic Driving Data. Transportation Research Board 96th Annual Meeting, Washington D.C., USA, 2017. pp. 1–12.
9. Lillicrap, T. P., J. J. Hunt, A. Pritzel, N. Heess, T. Erez, and, Y. Tassa. Continuous Control with Deep Reinforcement Learning. International Conference on Learning Representations, 2016.