Vehicle Trajectory Prediction based on Social Generative Adversarial Network for Self-Driving Car Applications

Li-Wei Kang
Department of Electrical Engineering
National Taiwan Normal University, Taipei, Taiwan
lwkang@ntnu.edu.tw

Chih-Chung Hsu
Department of Management Information Systems
National Pingtung University of Science and Technology
Pingtung, Taiwan
*cchsu@mail.npust.edu.tw

I-Shan Wang, Ting-Lei Liu, Shih-Yu Chen, and Chuan-Yu Chang
Department of Computer Science and Information Engineering
National Yunlin University of Science and Technology
Yunlin, Taiwan

***Abstract*—Self-driving or autonomous vehicles need to efficiently and continuously navigate in complex traffic environments by analyzing the surrounding scene, understanding the behavior of other traffic-agents, and predicting their future trajectories. The main goal is to plan a safe motion and reduce the reaction time for possibly imminent hazards. A critical and challenging problem considered in this paper is to explore the movement patterns of surrounding traffic-agents and accurately predict their future trajectories for helping the vehicle make reasonable decision. To solve the problem, a deep learning-based framework is proposed in this paper for predicting trajectories of autonomous vehicles. The key is to train a social GAN (generative adversarial network) deep model for prediction of vehicle trajectory. The presented experimental results have verified that the proposed social GAN-based approach outperforms the traditional Social LSTM (long short-term memory)-based method.**

Keywords-Self-driving, autonomous vehicles, vehicle trajectory, deep learning, generative adversarial network.

I. INTRODUCTION

Using autonomous vehicles in the near future would be expected to reduce the number of road accidents, improve road safety, and increase transport convenience [1], [2]. However, for safety and efficiency of traffics on roads, an autonomous vehicle should not only understand the current state of the nearby road-users, but also predict their future behaviors. One critical issue is the prediction of pedestrian behavior on roads, which has been investigated in the literature, such as [3], [4]. Another important problem is the prediction of the behavior of other vehicles on roads [2]. However, new challenges come from interdependencies among vehicles' behaviors, influence of traffic rules, and driving scenarios, and diversity of vehicle behaviors. Real limitations in exploring the surrounding scenario and the required computational resources to execute prediction algorithms also impose the difficulty on the problem. There were also several researches presented for solving the problem of vehicle behavior prediction in the literature, such as [5], [6].

Moreover, with the rapid development of deep learning theories and techniques [7], several vision-based applications, such as image classification/recognition [8]–[11], event detection [12], industry inspection [13], and image restoration [14]–[20], have achieved remarkable successes compared with traditional approaches. Therefore, several deep learning-based prediction frameworks for vehicle behaviors have been presented, such as [21]–[26]. This paper focuses on the prediction of vehicle trajectory and presents a deep learning-based framework to solve this problem.

II. OVERVIEW OF THE PROPOSED FRAMEWORK

The problem considered in this paper is to explore the movement patterns of surrounding vehicles and accurately predict their future trajectories for helping the vehicle make reasonable decision. In the proposed framework, we first apply the YOLO v3 deep object detector [27], trained on the MS-COCO dataset [28], for detection of vehicles. For tracking the detected vehicles, the deep SORT (simple online and realtime tracking) algorithm [29] is used. Based on the collected trajectory data generated from our traffic scenario dataset (with some examples shown in Fig. 1) using the deep SORT tracker, we train a social GAN (generative adversarial network) deep model [30] for prediction of vehicle trajectory.

Social GAN [30] was originally designed for the trajectory prediction of pedestrians. Social GAN presented a recurrent sequence-to-sequence model, observing motion histories and predicting future behavior by applying a pooling mechanism to aggregate information across people. Since our goal is to predict the trajectories of autonomous vehicles, we re-train the Social GAN by using the vehicle trajectories extracted from our collected dataset of traffic scenario images by using the above-mentioned YOLO v3 [27] and SORT [29] algorithms.

III. EXPERIMENTAL RESULTS

The proposed framework was implemented in Python programming language with Pytorch [31] on a personal computer equipped with Intel® Core™ i7-9700K CPU, 16 GB memory, and NVIDIA GeForce RTX 2080 Ti GPU. To train our Social GAN, we used our collected dataset and the Adam optimizer [32] with the initial learning rate set to 0.001 and the batch size set to 64. To evaluate the performance of the proposed model, we also re-trained the Social LSTM (long short-term memory) method (originally designed for human trajectory prediction) [33] using our traffic scenario dataset for comparisons with our framework. The two well-known

978-1-7281-9362-5/20/$31.00 ©2020 IEEE
DOI 10.1109/IS3C50286.2020.00133

Fig. 1. Some examples of our collected traffic scenario image dataset used for model training and testing.

metrics, ADE (average displacement error) and FDE (final displacement error), for trajectory prediction [23], were used for performance evaluation, as shown in Table I. Based on Table I, the proposed method achieves better results in prediction of final positions while achieving comparable results in prediction of average displacements. We also show some trajectory prediction results in Fig. 2.

TABLE I. PERFORMANCE EVALUATION FOR VEHICLE TRAJECTORY PREDICTION IN TERMS OF THE ADE AND FDE METRICES.

Metrics	**Evaluated Methods**	
	Social LSTM	***Proposed***
ADE	3.82	3.89
FDE	13.25	**6.61**

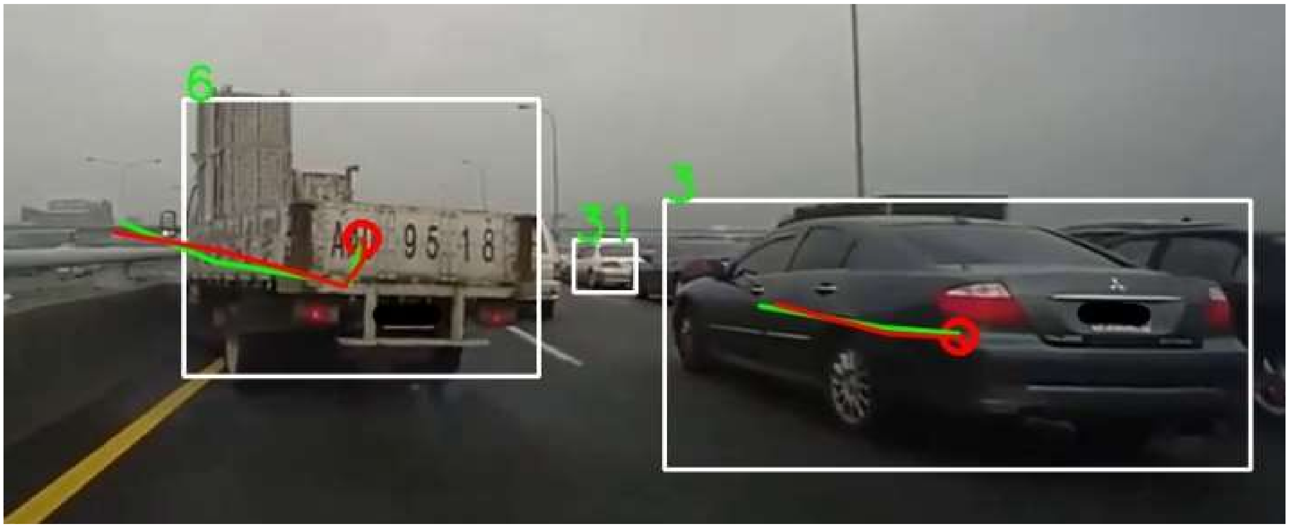

Fig. 2. The predicted trajectories obtained by the proposed method, denoted by red (the ground truths are denoted by green).

IV. Conclusions

In this paper, we have presented a social GAN (generative adversarial network)-based deep network for prediction of vehicle trajectory. By using the YOLO v3 deep object detector [27], trained on the MS-COCO dataset [28], and the deep SORT tracker [29], we extracted many vehicle trajectories from our collected dataset. Then the original Social GAN [30] designed for human movement prediction can be re-trained for prediction of vehicle trajectories. The presented experimental results have confirmed that our framework outperforms the re-trained Social LSTM model in vehicle trajectory prediction.

Acknowledgment

This work was supported in part by Ministry of Science and Technology, Taiwan, under the Grant MOST 108-2221-E-003-027-MY3.

References

[1] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 900–918, March 2020.

[2] S. Mozaffari, O. Y. Al-Jarrah, M. Dianati, P. Jennings, and A. Mouzakitis, "Deep learning-based vehicle behaviour prediction for autonomous driving applications: A review," *IEEE Trans. Intell. Transp. Syst.* (in Press).

[3] P. V. K. Borges, N. Conci, and A. Cavallaro, "Video-based human behavior understanding: A survey," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 23, no. 11, pp. 1993–2008, Nov. 2013.

[4] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, "Human motion trajectory prediction: a survey," *The Int. J. Robotics Research*, vol. 39, no. 8, pp. 895–935, 2020.

[5] M. S. Shirazi and B. T. Morris, "Looking at intersections: A survey of intersection monitoring, behavior and safety analysis of recent studies," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 4–24, Jan 2017.

[6] M. T. Brendan Tran Morris, "Understanding vehicular traffic behavior from video: a survey of unsupervised approaches," *J. Electronic Imaging*, vol. 22, no. 4, 2013.

[7] Y. LeCun, Y. Bengio, and G. E. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[8] C.-H. Yeh, M.-H. Lin, P.-C. Chang, and L.-W. Kang, "Enhanced visual attention-guided deep neural networks for image classification," *IEEE Access*, vol. 8, pp. 163447–163457, Sep. 2020.

[9] L.-W. Kang, K.-L. Chou, and R.-H. Fu, "Deep Learning-based weather image recognition," *Proc. IEEE Int. Sym. Computer, Consumer and Control*, Taichung, Taiwan, Dec. 2018.

[10] L.-W. Kang, T.-Z. Feng, and R.-H. Fu, "Inception network-based weather image classification with pre-filtering process," *Proc. Int. Computer Sym.*, 2018, pp. 368–375.

[11] S.-Y. Chen, L.-W. Kang, S.-C. Chen, and C.-Y. Chang, "Hyperspectral-guided drug classification via deep learning," *Proc. SPIE Optics and Photonics Int. Congress*, Pacifico Yokohama, Japan, Apr. 2019.

[12] L.-W. Kang, I.-S. Wang, K.-L. Chou, S.-Y. Chen, and C.-Y. Chang, "Image-based real-time fire detection using deep learning with data augmentation for vision-based surveillance applications," *Proc. IEEE Int. Conf. Advanced Video and Signal Based Surveillance*, Taipei, Taiwan, Sept. 2019.

[13] L.-W. Kang, Y.-T. Chen, W.-C. Jhong, and C.-Y. Hsu, "Deep learning-based identification of steel products," *Proc. IEEE Int. Conf. Intelligent Information Hiding and Multimedia Signal Processing*, Sendai, Japan, Nov. 2018.

[14] C.-Y. Lin, Z. Tao, A.-S. Xu, L.-W. Kang, and F. Akhyar, "Sequential dual attention network for rain streak removal in a single image," *IEEE Trans. Image Process.*, vol. 29, pp. 9250–9265, 2020.

[15] C.-H. Yeh, C.-H. Huang, and L.-W. Kang, "Multi-scale deep residual learning-based single image haze removal via image decomposition," *IEEE Trans. Image Process*., vol. 29, no. 1, pp. 3153–3167, Dec. 2020.

[16] C.-H. Yeh, C.-H. Huang, L.-W. Kang, and M.-H. Lin, "Single image dehazing via deep learning-based image restoration," *Proc. APSIPA Annual Summit and Conf*., Honolulu, Hawaii, USA, Nov. 2018.

[17] L.-W. Kang, C. Lo, and C.-H. Yeh, "Deep learning-based moiré pattern removal from a single image: A survey and comparative study," *Proc. IEEE Int. Conf. Consumer Electronics-Taiwan*, Taoyuan, Taiwan, Sep. 2020.

[18] M.-H. Lin, C.-H. Yeh, C.-H. Lin, and L.-W. Kang, "Deep multi-scale residual learning-based blocking artifacts reduction for compressed images," *Proc. IEEE Int. Conf. Artificial Intelligence Circuits and Systems*, Hsinchu, Taiwan, Mar. 2019.

[19] C.-H. Yeh, C.-H. Lin, M.-H. Lin, L.-W. Kang, C.-H. Huang, and M.-J. Chen, "Deep learning-based compressed image artifacts reduction

based on multi-scale image fusion," *Information Fusion*, vol. 67, pp. 195–207, March 2021.

[20] Z.-T. Zhang, C.-H. Yeh, and L.-W. Kang, and M.-H. Lin, "Efficient CTU-based intra frame coding for HEVC based on deep learning," *Proc. APSIPA Annual Summit and Conf.*, Kuala Lumpur, Malaysia, 2017.

[21] A. Zyner, S. Worrall, and E. Nebot, "Naturalistic driver intention and path prediction using recurrent neural networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1584–1594, 2020.

[22] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, "Attention based vehicle trajectory prediction," *IEEE Trans. Intelligent Vehicles* (in press).

[23] Y. Ma, X. Zhu, S. Zhang, R. Yang, W. Wang, and D. Manocha, "TrafficPredict: Trajectory prediction for heterogeneous traffic-agents," *Proc. AAAI Conf. Artificial Intelligence*, Honolulu, Hawaii, USA, Jan. 2019, pp. 6120–6127.

[24] N. Deo and M. M. Trivedi, "Convolutional social pooling for vehicle trajectory prediction," *Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops*, Salt Lake City, UT, USA, 2018, pp. 1468–1476.

[25] S. Dai, L. Li, and Z. Li, "Modeling vehicle interactions via modified LSTM models for trajectory prediction," *IEEE Access*, vol. 7, pp. 38 287–38296, 2019.

[26] W. Luo, B. Yang, and R. Urtasun, "Fast and furious: Real time endto-end 3d detection, tracking and motion forecasting with a single convolutional net," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, June 2018.

[27] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv:1804.02767*, 2018.

[28] T.-Y. Lin, M. Maire, S. J. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," *Proc. European Conf. Computer Vision*, Zurich, Switzerland, Sept. 2014, pp. 740–755.

[29] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," *Proc. IEEE Int. Conf. Image Processing,* Beijing, China, Sept. 2017, pp. 3645–3649.

[30] A. Gupta, J. Johnson, Li F.-F., S. Savarese, and A. Alahi, "Social GAN: Socially acceptable trajectories with generative adversarial networks," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, June 2018, pp. 2255–2264.

[31] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," *Proc. Annual Conf. Neural Information Processing Systems*, Vancouver, BC, Canada, Dec. 2019, pp. 8024–8035.

[32] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *Proc. Int. Conf. Learning Representations*, May 2015.

[33] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, Li F.-F., and S. Savarese, "Social LSTM: Human trajectory prediction in crowded spaces, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, June 2016, pp. 961–971.