Discovering Human-like Driving Strategies with Learning and Optimization Approaches

Erik Dovgan*, Ivan Bratko**, Jaka Sodnik***, ****

* Department of Intelligent Systems, Jožef Stefan Institute, Ljubljana, Slovenia

** Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia

*** Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

**** NERVteh, raziskave in razvoj, d.o.o., Trzin, Slovenia

erik.dovgan@ijs.si, bratko@fri.uni-lj.si, jaka.sodnik@fe.uni-lj.si

***Abstract*— Human driving strategies can be obtained in various ways. This paper presents two artificial intelligence approaches: a learning approach that builds a human driving model based on driving data, and an optimization approach that searches for driving strategies by taking into account existing human driving models. The results show that the optimization approach can optimize several objectives in addition to human likeness thus producing a set of driving strategies with various tradeoffs between the objectives, while the learning approach aims at exactly reproducing the human driving, which results in a single driving strategy.**

I. INTRODUCTION

Autonomous vehicle driving has recently been investigated by several automotive and other companies. The main task of autonomous driving consists of defining driving strategies that meet various objectives and constraints, such as safety, low energy consumption, drivability etc. To this end, several aspects have to be taken into account, such as the behavior of other human-driven vehicles on the route. The driving strategies are extensively evaluated in the driving simulators before the deployment in real vehicles. To make the driving environment in the simulators as realistic as possible, the traffic, i.e., the behavior of human drivers has to be appropriately replicated. Therefore, an effective replication of human driving is of key importance for evaluation the driving strategies in various simulated real-life situations. The obtained driving strategies that replicate human driving and/or optimize other objectives can be used for both driving of autonomous vehicles and driving of human-like traffic vehicles.

The presented work on discovery of human-like driving strategies consists of two approaches. The optimization approach focuses on non-critical driving situations, such as car following, free driving, and overtaking. The existing research in this field mainly focuses on the single-objective optimization and the discovery of driving strategies that mimic human driving as much as possible. However, besides human likeness, several other objectives should be taken into account, such as the traveling time and the fuel consumption. The presented paper proposes an approach that takes these additional objectives into account.

On the other hand, the learning approach is applied on the critical driving situations, such as collision avoidance. Existing approaches for learning human driving mainly focus on non-critical driving situations or on specific critical situations. The proposed approach, on the other hand, aims at handling various critical situations by implementing a generic representation of driving situations and using it as an input for the deep neural network.

The paper is further organized as follows. The related work is presented in Section II. Section III presents the optimization and learning approaches for discovering driving strategies. The numerical experiments are presented in Section IV. Finally, Section V concludes the paper with ideas for future work.

II. RELATED WORK

Human driving can be mimicked by either applying appropriate human driving models or learning the driving models. In the first case, the predefined driving models in the form of sets of equations are optimized to drive the vehicle as similar as possible to humans. In the second case, the appropriate learning algorithms construct the driving models without any predefined model structure.

Several human driving models have been proposed by the researchers. These models typically focus on specific driving situations, such as car following and overtaking. Car following models are applied when the vehicle follows the preceding vehicle on the same lane. To that end, they control the vehicle by responding to the stimulus, i.e., actions of the preceding vehicle. For example, when the preceding vehicle accelerates/decelerates, the following vehicle has also to accelerate/decelerate. However, this response is not done instantaneously, but tries to imitate the response of human drivers. The best-know stimulus-response model is the Gazis-Herman-Rothery (GHR) model, which specifies the stimulus as the relative velocity between the vehicles. Additional parameters that have been included in the model comprise the distance to the preceding vehicle, the velocity of the vehicles, different parameters for acceleration and deceleration phases etc. [3,13]. The linear model defines the vehicle acceleration as a linear combination of the desired following distance, the vehicle velocity, the relative distance to the preceding vehicle, the relative velocity to the preceding vehicle, and the reaction time [26]. The safety distance model drives the vehicle at a safe following distance by manipulating basic Newtonian equations of motion [10]. The optimal velocity model achieves the optimal velocity based on the distance to the preceding vehicle and the velocity difference to the preceding vehicle [16]. The psychophysical model defines a set of action points, i.e., thresholds, which are used to trigger actions when the selected stimuli (distance, velocity difference etc.) fall beyond their threshold values [22]. The cellular automaton model describes the traffic as a stochastic discrete automaton model. The goal is to not have two vehicles in the same cell at the same time, i.e., to avoid collisions [17].

Other types of driving models that have been widely developed are the lane changing and overtaking models. These models determine when the lane should be changed or when the overtaking should start. This is typically determined in two steps. In the first step, the desire to start lane changing or overtaking is determined, while in the second step, the gap on the target lane is checked. If both conditions are satisfied, the vehicle starts to change the lane or overtake. Examples of these models can be found in [6,7,12]. However, to control the vehicle during the overtaking/lane changing, these models need to be coupled with the car following models. As a consequence, both types of models can be simultaneously optimized to obtain better driving strategies. Examples of such a simultaneous optimization can be found in [1,18,19].

Driving strategies imitating human driving can be also obtained with learning approaches. These approaches typically focus on specific driving situations, such as car following, overtaking, etc. In contrast to the optimization approaches, where the models are predefined and include intuitive relations among the input variables, the learning models are not predefined and are not intuitive, since the modeled relations are hard to explain and understand. The learning approaches mainly differ in the learning algorithms that are used to build the driving models. Examples of car following use conditional linear Gauss model [15], Gaussian mixture model [2], PrARX model [14], and Support Vector Regression [23]. To detect maneuvers in crossroads, Gaussian process regression can be used [20]. To drive in the city and on highway, artificial neural networks have been applied [25]. Deep neural networks have recently also been applied to model human driving. These models aim at solving end-to-end driving task, which consists of learning to drive by taking as the input images from cameras. This differs significantly from the previously listed approaches, which preprocess raw data, reconstruct the scene and only afterwards apply the learning approaches. Examples of such end-to-end approaches that use deep neural networks can be found in [5,8,11,24].

The existing optimization methods focus on modeling human driving behavior and ignore other objectives that are also important when driving a vehicle, such as short traveling time and low fuel consumption. This paper presents an algorithm that considers both aspects, since it implements human driving models to obtain human-like driving, and optimizes the traveling time and the fuel consumption. This paper also presents a novel approach for applying deep neural networks for learning human driving. This approach enhances the existing approaches by introducing a generic representation of the driving situation that is used as the input to the deep neural network.

III. Discovering Human-like Driving Strategies

The discovery of human-like driving strategies employs two approaches. The first approach is based on the optimization of parameter values of the predefined human driving models and aims at optimizing the traveling time and the fuel consumption. The second approach, on the other hand, aims at reproducing human driving with machine learning algorithms and does not take additional objectives into account. In addition, the optimization approach focuses on non-critical driving situations, while the learning approach aims at driving the vehicle in critical driving situations. The main difference between the two approaches is that when using the learning approach, objectives such as the traveling time and the fuel consumption cannot be optimized, since leaning only aims at replicating the given human driving. That is why two independent tasks have been defined, one optimization and one learning, and appropriate approaches have been developed for each task independently.

A. *Optimization Approach for Discovering Human-like Driving Strategies*

The optimization of human-like driving strategies for handling non-critical driving situations consists of a two-level multiobjective optimization algorithm that searches for driving strategies, which simultaneously mimics human driving behavior and minimizes the traveling time and the fuel consumption. On the lower level, a set of human driving models are used to determine the control actions that are applied during vehicle driving. The parameters of these models are tuned with the upper-level algorithm, i.e., a multiobjective evolutionary algorithm, which searches for a set of driving strategies that are nondominated with respect to the traveling time and the fuel consumption.

The lower level consists of a set of mathematical models that aim to reproduce human driving in the following set of driving situations: car following, free driving, emergency deceleration, and overtaking. These models determine the vehicle's longitudinal acceleration and the moment when the overtaking should start.

The car following model is based on the GHR model [3], which determines the acceleration based on the current velocity, the distance to the preceding vehicle and the velocity difference with respect to the preceding vehicle, as shown in (1). For more details, see [3].

$$a_t = \begin{cases} k_1 {v_t}^{k_2} {\Delta s_t}^{k_3} \Delta v_{t-\tau} & ; \Delta v_{t-\tau} > 0 \\ k_4 {v_t}^{k_5} {\Delta s_t}^{k_6} \Delta v_{t-\tau} & ; \Delta v_{t-\tau} < 0 \\ 0 & ; \Delta v_{t-\tau} = 0 \end{cases} \quad (1)$$

The free driving model and the emergency deceleration model aim at achieving the target velocity with an acceleration that is calculated with a linear function of the current velocity. The overtaking models calculate the driver's utility to pass and the minimal gap as functions of the current velocity, velocity of the preceding vehicle, velocity of the vehicle on the other lane, and gaps on both lanes, as shown in (2) and (3). When the utility is greater than zero and the available gap is larger than the minimal

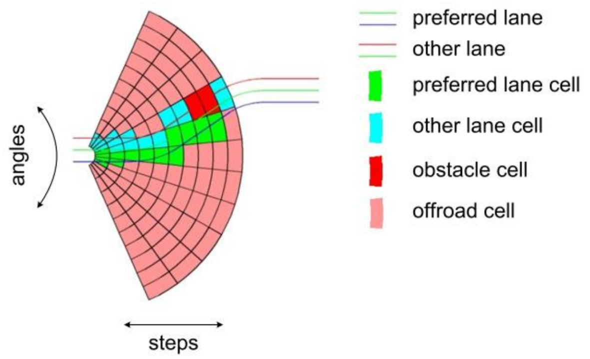

Figure 1. Generic representation of driving situation consisting of the space in front of the vehicle that is partitioned into cells

gap, the vehicle starts to overtake. For more details, see [7].

$$u = k_7 + k_8 \Delta v_{\text{target}} + k_9 \Delta s \qquad (2)$$

$$g = k_{10} + k_{11} v + k_{12} v_{\text{preceding}} + k_{13} v_{\text{opposite}} \qquad (3)$$

The upper level of the multiobjective optimization algorithm consists of a multiobjective evolutionary algorithm that is based on DEMO [21] and NSGA-II [4]. This algorithm optimizes the traveling time and the fuel consumption by tuning the parameters of the lower-level human driving models, and returns a set of nondominated driving strategies. To this end, the algorithm operates with a set of solutions called population, where each solution stores the values of all the parameters of human driving models from the lower-level algorithm. These solutions are randomly initialized and then improved in several iterations called generations by employing evolutionary operators, i.e., selection, crossover and mutation. More precisely, the algorithm uses the scheme DE/rand/1/bin [21] to create new solutions from the existing solutions. After each generation, the best solutions are selected to remain in the population thus maintaining a constant population size. For this purpose, multiobjective mechanisms from NSGA-II are used. Finally, a set of nondominated solutions is returned.

B. *Learning Approach for Discovering Human-like Driving Strategies*

Human-like driving strategies for critical situations are learned with the deep neural network algorithm [9]. For the input of the network, a generic representation of the driving situation has been developed. This representation discretizes the space in front of the vehicle into cells, where each cell stores the information regarding the presence of road, obstacles, etc. More precisely, only a limited space in front of the vehicle is taken into account, i.e., the space within an angle range with respect to the vehicle heading and a given length ahead of the vehicle. This space is partitioned into cells with respect to various angles and distances from the vehicle. Each cell stores the information regarding the road lane (preferred lane or other lane), road position (on- or off- road), and presence of obstacles such as other vehicles. An example of the space partitioning can be seen in Fig. 1.

The deep neural network algorithm builds the human driving model in the form of a deep network that consists

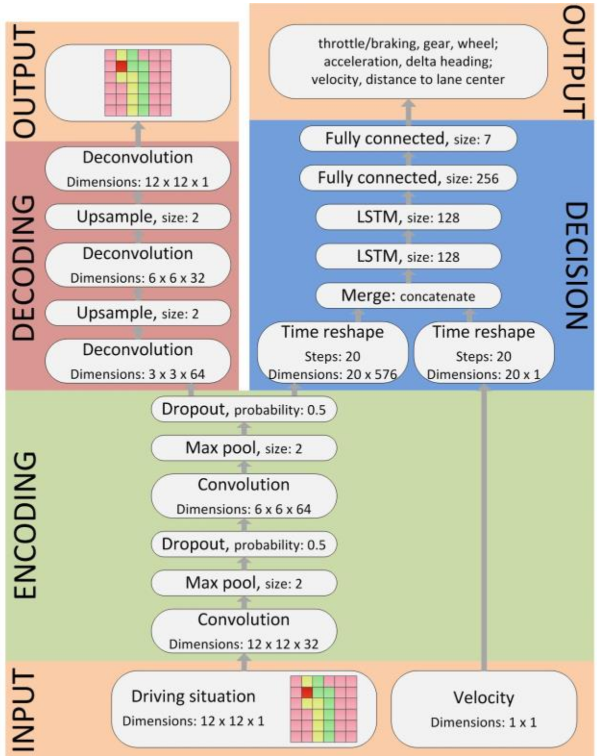

Figure 2. The architecture of the neural network. Left-hand side: the autoencoding model consisting of encoding and decoding layers. Right-hand side: the final model consisting of the encoding layers trained with autoencoding, and decision layers

of several layers as shown in Fig. 2. The network input includes the driving situation and the vehicle velocity. The network returns seven data that can be used to control the vehicle. These control data are combined in three groups {a, b, c} as follows: (a) throttle/braking percentage, gear and wheel angle, (b) acceleration and change in vehicle heading, and (c) velocity and distance to the center of the right lane. Only one group among {a, b, c} is used to control the vehicle. Nevertheless, all three groups were evaluated to determine the best group of control data.

The developed network consists of several encoding and decision layers [9] as shown in Fig. 2. The encoding layers include convolution, max pool and dropout layers, and aim at learning important features from the situation data. The decision layers include long short-term memory (LSTM) and fully connected layers, and aim at discovering temporal patters by processing the encoding data within a sliding time window. These layers are trained in two steps. In the first step, the encoding layers are trained by employing the autoencoding technique [9], while in the second step, the decision layers are coupled with the pretrained encoding layers and only the decision layers are trained. The autoencoding technique adds "mirror" layers to the encoding layers in order to produce the output that is the same as the input. Therefore, the convolution layers are mirrored with deconvolution layers, and the max pool layers are mirrored with the upsample layers. Finally, the algorithm returns the deep network that consists of the encoding and decision layers.

IV. EXPERIMENTS AND RESULTS

The proposed approaches were tested in two driving scenarios. To test the optimization approach, a two-lane

TABLE I.
SUMMARY OF THE EVALUATION OF DEEP NEURAL NETWORK

observed set of data	statistic	throttle/ braking [%]	gear	wheel [rad]	acceleration [m/s^2]	delta heading [delta rad]	velocity [m/s]	distance to lane center [m]
evaluation set	average	0.16	2.29	0.15	0.07	0.00	8.03	-0.89
	σ	0.28	1.17	0.76	1.08	0.06	5.86	1.75
	min	-1.00	-1.00	-4.34	-17.53	-0.67	-2.17	-8.28
	max	1.00	6.00	7.85	5.36	1.29	27.99	5.04
predicted data	root mean square error	0.22	0.97	0.51	0.86	0.04	0.79	0.20

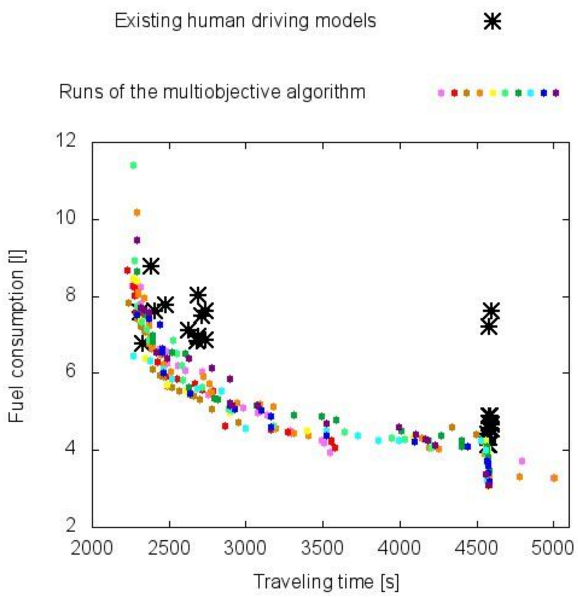

Figure 3. Comparison of driving strategies obtained with the multiobjective algorithm and existing human driving models

rural route was used and the results were compared to the existing human driving models. The learning approach was, on the other hand, tested on one critical situation in the urban environment, where a moving vehicle had to be avoided.

A. Experimental Setup

Both tested scenarios included traffic vehicles. In addition, the critical scenario included pedestrians that crossed the road on zebra crossing. The length of the rural route was 46 km, while the length of the urban route was around one kilometer. The critical situation in the urban environment consisted of a truck moving backwards from a parking area to the main street. This parking area was surrounded by a high wall that prevented to see the truck until the driven vehicle was close to the truck.

The learning algorithm requires a set of driving data, which is used in the learning and evaluation phases. To this end, driving data of 200 drivers were recorded with a simulator based on the SCANeR Studio[1], which was provided by the NERVteh company[2].

The optimization algorithm was tested with the following parameter setup: population size: 50, number of generations: 50, crossover probability: 0.9, scaling factor: 0.5. Since it is a stochastic algorithm, it was run 10 times to obtain the driving strategies.

To evaluate the learning algorithm, the following parameter values were used: number of epochs: 200, optimizer: simple gradient descent, learning rate: 0.1, learning rate decay: 0.96, decay step: 1000, data split: 60 % for model training and 40 % for model evaluation. The same configuration was used to train the main network and for the autoencoding technique.

B. Experimental Results

Fig. 3 shows the driving strategies for non-critical driving situations in the objective space, obtained with 10 runs of the multiobjective optimization algorithm. In addition, this figure also shows the driving strategies obtained with the existing approaches that were presented in [1,6,7]. These results show that the existing approaches focus either on the car following without overtaking (the results on the right-hand side of the figure and with long traveling time) or on overtaking (the results on the left-hand side of the figure and with short traveling time). However, they fail to discover driving strategies with compromises between the two objectives. On the other hand, the multiobjective optimization algorithm finds also a large set of driving strategies with various tradeoffs between the objectives. Such a set of driving strategies enables the user to have a better overview of all the possible driving strategies and to select the preferred one without limiting the tradeoff between objectives. In addition, it also enables to deploy a large set of driving strategies with various tradeoffs between objectives in a driving simulation environment.

The deep neural network has been evaluated with three sets of control actions as described in Section III.B. The results are shown in Table 1 and Fig. 4. This table shows that the standard deviation in the evaluation set is similar to the root mean square error in the predicted data when the vehicle is controlled with {throttle/braking percentage, gear, wheel angle} or {acceleration, change in vehicle heading}. On the other hand, the root mean square error is significantly lower than the standard deviation when the vehicle is controlled with {velocity, distance to the center of the right lane}. Consequently, when using deep neural networks with the proposed generic representation of the driving situation, the preferred actions for controlling the vehicle are {velocity, distance to the center of the right lane}. This is also confirmed by Fig. 4 that shows the comparison between the predicted and the actual data on

[1] http://www.oktal.fr/en/automotive/range-of-simulators/software

[2] https://www.nerv-teh.com/

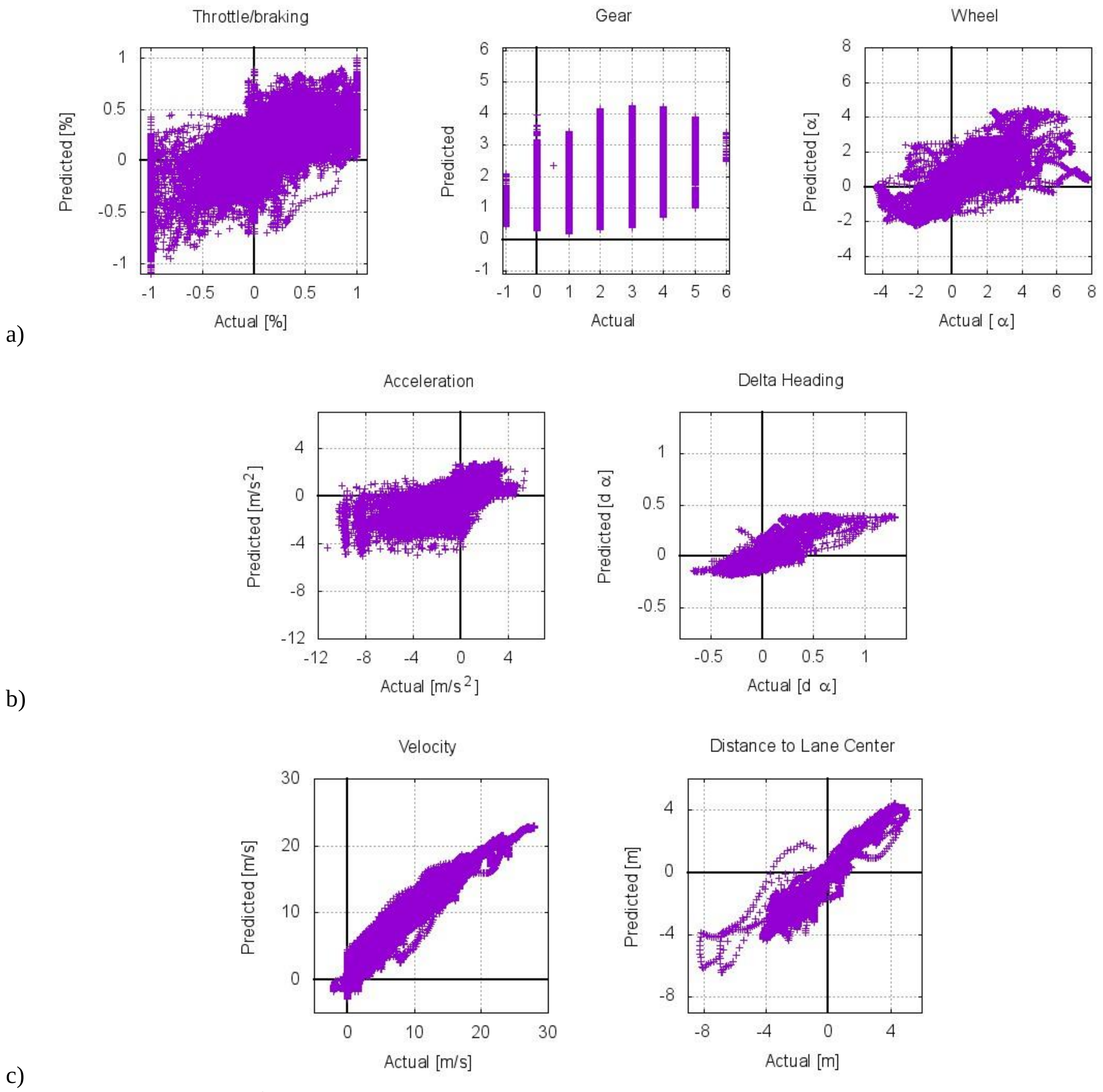

Figure 4. Comparison of predicted and actual values, obtained by evaluating deep neural network driving model that predicted: a) throttle/braking percentage, gear, and wheel angle; b) acceleration and change in vehicle heading; c) velocity and distance to the center of the right lane

the evaluation set for all seven control actions. This figure shows that among the predicted control actions, the closest to the ideal values represented by the line $y = x$ are the velocity and the distance to the center of the right lane. On the other hand, wheel, acceleration and change in vehicle heading are mostly underestimated, while throttle/braking percentage and gear are predicted almost at random.

V. CONCLUSIONS

We designed and evaluated two approaches for discovering human-like driving strategies. The learning approach uses the deep neural network algorithm to build driving strategies for critical driving situations. In addition, the optimization approach implements the multiobjective optimization algorithm that discovers driving strategies by taking into account additional objectives, i.e., the traveling time and the fuel consumption, and is used to handle non-critical driving situations. The results show that the multiobjective optimization algorithm is able to find good driving strategies with various tradeoffs between the objectives. In addition, the results also show that deep neural network is able to replicate human driving when the vehicle is controlled with velocity and distance to the center of the right lane. In our future work, we will evaluate the deep neural network in additional critical driving situations. In addition, we will compare its results with existing learning approaches.

ACKNOWLEDGMENT

The authors acknowledge the project ("Multiobjective discovery of driving strategies for autonomous vehicles," funding No. Z2-7581) was financially supported by the Slovenian Research Agency. The authors also acknowledge the financial support from the Slovenian Research Agency (research core funding No. P2-0209).

This work was also in part funded by NERVteh, raziskave in razvoj, d.o.o., and is part of a project that has received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement no. 692286.

References

[1] K. I. Ahmed, *Modeling Drivers Acceleration and Lane Changing Behavior*, Ph.D. Dissertation, Cambridge: Massachusetts Institute of Technology, 1999.

[2] P. Angkititrakul, C. Miyajima, and K. Takeda, “Modeling and adaptation of stochastic driver-behavior model with application to car following,” *Proceedings of the IEEE Intelligent Vehicles Symposium IV*, pp. 814–819, 2011.

[3] R. E. Chandler, R. Herman, and E. W. Montroll, “Traffic dynamics: studies in car following,” *Operations Research*, vol. 6, no. 2, pp. 165–184, 1958.

[4] K. Deb, A. Pratap, S. Agrawal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[5] H. M. Eraqi, M. N. Moustafa, and J. Honer, “End-to-end deep learning for steering autonomous vehicles considering temporal dependencies,” *CoRR*, vol. abs/1710.03804, 2017.

[6] H. Farah, S. Bekhor, A. Polus, and T. Toledo, “A passing gap acceptance model for two-lane rural highways,” *Transportmetrica*, vol. 5, no. 3, pp. 159–172, 2009.

[7] H. Farah and T. Toledo, “Passing behavior on two-lane highways,” *Transportation Research Part F*, vol. 13, no. 6, pp. 355–364, 2010.

[8] T. Fernando, S. Denman, S. Sridharan, and C. Fookes, “Going deeper: Autonomous steering with neural memory networks,” in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 214–221, 2017.

[9] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge: Massachusetts Institute of Technology, 2016.

[10] M. R. Hafner, D. Cunningham, L. Caminiti, and D. Del Vecchio, “Cooperative collision avoidance at intersections: algorithms and experiments,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1162–1175, 2013.

[11] Y. Hou, S. Hornauer, and K. Zipser, “Fast recurrent fully convolutional networks for direct perception in autonomous driving,” *CoRR*, vol. abs/1711.06459, 2017.

[12] A. Kusuma, R. Liu, and F. Montgomery, “Gap acceptance behavior in motorway weaving sections,” in *Proceedings of the Eastern Asia Society for Transportation Studies*, vol. 9, 2013.

[13] Y. Li and D. Sun, “Microscopic car-following model for the traffic flow: the state of the art,” *Journal of Control Theory and Applications*, vol. 10, no. 2, pp. 133–143, 2012.

[14] K. Mikami, H. Okuda, S. Taguchi, Y. Tazaki, and T. Suzuki, “Model predictive assisting control of vehicle following task based on driver model,” in *Proceedings of the IEEE Conference on Control Technology and Applications*, pp. 890–895, 2010.

[15] D. Moser, R. Schmied, H. Waschl, and L. del Re, “Flexible spacing adaptive cruise control using stochastic model predictive control,” *IEEE Transactions on Control Systems Technology*, vol. 26, no. 1, pp. 114–127, 2018.

[16] D. Sun, Y. Li, and C. Tian, “Car-following model based on the information of multiple ahead and velocity difference,” *System Engineering Theory and Practice*, vol. 30, no. 7, pp. 1326–1332, 2010.

[17] T. Toledo, “Driving behaviour: models and challenges,” *Transport Reviews*, vol. 27, no. 1, pp. 65–84, 2007.

[18] T. Toledo, *Integrated Driving Behaviour Modelling*, Ph.D. Dissertation, Cambridge: Massachusetts Institute of Technology, 2002.

[19] T. Toledo, H. N. Koutsopoulos, and M. E. Ben-Akiva, “Modeling integrated lane-changing behavior,” *Transportation Research Record*, vol. 1857, pp. 30–38, 2003.

[20] Q. Tran and J. Firl, “Modelling of traffic situations at urban intersections with probabilistic non-parametric regression,” in *Proceedings of the IEEE Intelligent Vehicles Symposium IV*, pp. 334–339, 2013.

[21] T. Tušar and B. Filipič, “Differential evolution versus genetic algorithms in multiobjective optimization,” in *Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization EMO*, pp. 257–271, 2007.

[22] W. Wang, W. Zhang, D. Lia, K. Hirahara, and K. Ikeuchi, “Improved action point model in traffic flow based on driver’s cognitive mechanism,” in *Proceedings of the 2014 IEEE Intelligent Vehicles Symposium*, pp. 447–452, 2004.

[23] D. Wei and H. Liu, “Analysis of asymmetric driving behavior using a self-learning approach,” *Transportation Research Part B: Methodological*, vol. 47, pp. 1–14, 2013.

[24] H. Xu, Y. Gao, F. Yu, and T. Darrell, “End-to-end learning of driving models from large-scale video datasets,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3530–3538, 2017.

[25] L. Xu, J. Hu, H. Jiang, and W. Meng, “Establishing style-oriented driver models by imitating human driving behaviors,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2522–2530, 2015.

[26] T. H. Yang and C. W. Zu, “Linear dynamic car-following model,” in *Proceedings of the 5th World Congress on Intelligent Control and Automation*, vol. 6, pp. 5212–5216, 2004.