2018 IEEE Intelligent Vehicles Symposium (IV)
Changshu, Suzhou, China, June 26-30, 2018

Optimization of the Braking Strategy for an Emergency Braking System by the Application of Machine Learning

Markus Schratter, Sabine Amler, Paul Daman, and Daniel Watzenig, *Senior Member, IEEE*

***Abstract*— This paper explores a methodology whereby accident data is directly used to develop a braking strategy for an autonomous emergency braking system. Future vehicles will be equipped with additional technologies. Detailed information about accidents or critical situations can be recorded. If enough recorded data from critical situations are available, such data could be used to improve Active Safety Systems. In our approach, we do not model the behavior of pedestrians or drivers. The idea is to use the capability of machine learning to get the behaviors out of traffic data. Machine learning is used to derive the function design for an emergency braking system for pedestrians. Generated traffic scenarios are used to review the methodology. Random Forests and Neural Networks are used for the function designs and the learned function designs are compared with a reference implementation.**

***Index Terms*— Collision Avoidance, Automated Vehicles, Active Safety, Machine Learning, Development Process**

INTRODUCTION

Available Active Safety Systems already address the most common accidents, such as rear-end collisions or accidents with pedestrians crossing the road. Therefore, in the development process such accidents are analyzed and necessary strategies are derived. This approach works very well and achieves a high coverage of the most common accidents. In fact, a detailed analysis of accident data by the German Insurers Accident Research [1] has shown the 26 most common combinations of types and causes of accidents which account for 50% of all accidents. However, the remaining 50% of accidents are caused by 5287 combinations of types and causes of accidents. Significantly increasing the coverage of accidents requires too much effort in the development of individual Active Safety Systems.

A manageable system requires new approaches that can address a larger variation of critical situations. Using the technology built into Highly Automated Driving (HAD) vehicles, the vehicle itself can record the accident sequence. Based on this recorded data, we investigated how this data can be used to learn the behavior needed for a pedestrian protection system. In our approach, we do not want to model the behavior of pedestrians or drivers. The idea is to use the capability of machine learning to get the behavior out of traffic data, and derive the function design for the Active Safety System from the data. Thereby, we use machine learning to predict the request for a brake intervention. Only the system limitations of the Active Safety System need to be modeled during data preprocessing for the supervised learning task. Until now, there are no recorded accidents available with the required data quality for this purpose. Therefore, we use scenarios generated from an effectiveness analysis to evaluate the methodology.

M. Schratter is with Virtual Vehicle Research Center, 8010, Graz, Austria (markus.schratter@v2c2.at).
S. Amler and P. Daman are with BMW Group, 80788, Munich, Germany.
D. Watzenig is with Virtual Vehicle Research Center and also with Institute of Automation and Control, Graz University of Technology, 8010, Graz, Austria

I. Related Work

Currently machine learning is widely used in the context of autonomous driving, especially deep learning for computer vision tasks but also for the complete driving tasks, like the end-to-end learning approach showed in [2]. End-to-end learning shows how a vehicle can autonomously drive with a deep neural network, but also clearly shows the fundamental limitations of the method, like handling of complex driving tasks including traffic rules. A different architecture with a combination of programmed algorithms and neural networks for defined tasks (e.g. object detection, free space detection, path planning, etc.) trained with real world and generated data is illustrated in [3].

Machine learning is used to predict further driving trajectories of the vehicle in [4] and [5]. In [6] the path prediction of pedestrians is described by using machine learning. A Markov decision process in combination with deep reinforcement learning is investigated in [7], where the function design is learned with a reward function on generated scenarios.

In our approach, we learn the braking strategy and not just partial aspects of it, like predicting the driver or pedestrian behavior.

II. Motivation

A. Problem: Accident Collection

Active Safety Systems already address a high proportion of the most common accidents. Different groups of experts analyze accidents and derive requirements, and they also evaluate what is technically reasonable and feasible. In general therefore, two types of accident databases are used:

1) Quantitative accident databases (e.g. DESTATIS [8]): High amount, less detailed information, e.g. collision speed, collision point, time, etc.

+ Gives an overview about common accidents
- No information about the sequence of events

2) Qualitative accident databases (e.g. GIDAS Pre-Crash-Matrix [9]): Measurement and analyzes of the accident site, detailed specific reconstruction of the accident

+ Detailed information about the sequence of events
- Extensive preprocessing and reconstruction

Quantitative accident data illustrates quite well, which types of traffic accidents occur, but they do not deliver accurate information about the sequence of the accident. The collection of qualitative accident data is an elaborate process and is mostly related to a regional area. Global issues cannot fully be answered, because of the wide range of varying traffic rules and different driver behaviors in individual countries. Both data sources have the disadvantage of being collected after an accident has occurred. They are partially incomplete and contain many uncertainties (reconstruction).

B. Complexity and Variation of Accidents

Available Active Safety Systems already address the most common accidents with the corresponding common accident causes. Therefore an analysis is made to find out which factors are necessary to address a type of accident (e.g. position of the pedestrian). This approach works very well and a high coverage in the field is reached, as described in [10]. A detailed analysis of accident data has shown [1] that already 50% of all accidents are covered by the 26 most combinations of types and causes of accidents. However, the remaining 50% of accidents are caused by 5287 possible combinations of types and causes of accidents. A further significant increase in the addressing of accidents leads to an exponential increase in the effort that is necessary for developing Active Safety Systems. For a manageable system, innovative approaches are needed, which are able to address a larger variation of critical situations.

C. Highly Automated Driving Technology

To bring Highly Automated Driving functions on the road, the vehicle needs to understand the entire traffic situation in order to control the driving task by itself. Therefore, a lot of technologies need to be integrated into the vehicles. In [11] it is well described which technologies are needed to bring the functionality on the road. A more detailed description about the HAD technology can be found in the paper. Available technologies from a functional point and which are needed for the approach are the following:

- 360° Redundant environment recognition,
- High-precision Digital Maps,
- Driver monitoring,
- Backend communication,
- High available actuators.

With these technologies, a more accurate recognition of the environment and a more detailed interpretation of the traffic situation is achievable. Additionally, if a critical situation occurs, the built-in technology in HAD vehicles is able to record the accident sequence of events through the vehicle itself. This requires strict compliance to the legal framework for data protection.

III. System Design

The idea of the approach is to use the capability of machine learning to get the driver and pedestrian behaviors out of the traffic data and derive the function design for an Active Safety System. Machine learning is widely used to handle and learn the uncertainties in models, and to describe the behaviors of complex systems. In our approach, we do not want to model the behavior of pedestrians or drivers. The behavior should be modeled out of traffic data with the usage of machine learning. Our approach is split into four parts, which is shown schematically in Fig. 1. The following paragraphs describe the individual work items of the methodology evaluation.

A. Approach

1) Data Collection and Preprocessing

Traffic data with critical and non-critical situations are needed. Additionally, with existing reconstructed accidents, HAD vehicles can be used for a detailed recording of accidents. With the built-in technology in HAD vehicles, it is possible to record the accident sequence of events through the vehicle itself. The recorded data allows a detailed analysis of the accidents and to derive critical factors of the situation. Furthermore, it is also possible to use the recorded data directly for the function development of Active Safety Systems. Therefore, following information from HAD vehicles can be used:

1. 360° environment recognition which captures all relevant road users.
2. Detailed representation of the road topology and the mapping of other road users on the road.
3. Driver monitoring to capture the drivers behavior during critical situations.
4. Backend communication to the car manufacturer to exchange data from recorded critical situations.

Currently no recorded real accident data is available for our method analysis in the required quality. Therefore generated scenarios are used. In our approach the dataset of traffic scenarios includes every time step of the course of the vehicle and the pedestrian. Additionally, the accident severity is available if an accident occurs. Based on this information, a dataset is generated, which describes the situation for each time step in detail. This is the starting point for the approach.

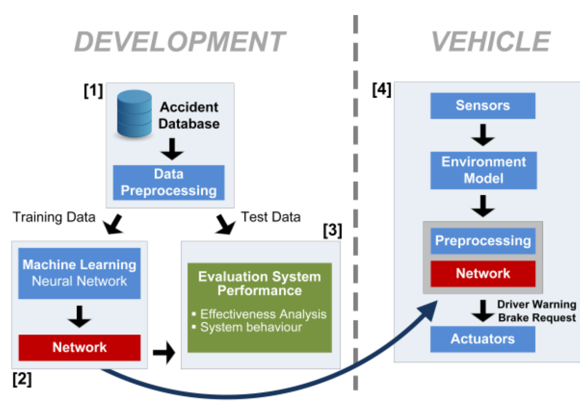

Figure 1. Schematic overview of the approach. (1) Data collection and preprocessing. (2) Machine learning of the function design. (3) Evaluation of the system performance. (4) Transferring of the function design into the vehicle.

In the next step, all time steps for every scenario need to be labeled for the supervised learning of the function design. For the function design we would like to know when a braking intervention is necessary to avoid an accident. Therefore, each dataset is classified in terms of the avoidance potential. The avoidance potential depends on the system limitations, such as the braking performance and including the road conditions, signal propagation delays, sensor parameters and functional requirements. These parameters need to be considered in the function design. Uncertainties in the movement of the pedestrian or the driver are not considered, these influences should be determined by the machine learning algorithm. The labeled dataset is used in the next step as training and test data. The following pseudo code shows how the dataset is labeled.

Pseudo code: Data Preprocessing / Labeling

```
foreach scenario:
  get_all_data_from_screnario()
  foreach time_step:
    features = get_data_current_previous_time_steps()
    time_to_brake = get_current_brake_time()
    predict_pedestrian_position_at_TTC_0s()
    if (time_to_brake > time_to_collision) and
      (collision_predicted == true) then
          avoidable = 0
    else
          avoidable = 1
    end
    add_dataset(features, avoidable)
  end
end
```

2) Machine Learning of the Function Design

At the beginning of our research [12] we only used random forests as machine learning algorithm. In this paper, we additionally investigated the usage of Neural Networks to predict the probability of a required brake trigger. The learned network delivers a probability for a brake request. The system performance can be fine-tuned with a defined threshold for the probability of the brake request.

One method to improve the quality of machine learning algorithms is through enrichment of data by additionally added knowledge based on the original data. Here, we generate from the existing data the time-to-collision (TTC), and the predicted lateral offset at the estimated collision point. The TTC results from the distance and the relative speed between the pedestrian and the vehicle. With this additional redundant data it is not guaranteed that the algorithm performs better. Tab. 1 shows the list of used features. The feature set additionally includes a time series for some features. Feature 16 and 17 are only used in the extended feature set.

3) Evaluation of the System Performance

To evaluate the system performance of the machine learning algorithms, we compare the performance to a reference system. The behavior of the reference system corresponds to the AEB pedestrian system in series production, with some slight variations in some system parameters. The following methods of measurements are used in detail to compare the algorithms:

- Analysis of the effectiveness [13]

The effectiveness of the algorithm shows how many accidents the algorithm avoids and how high the average speed reduction of the algorithm is.

- Receiver-Operating-Characteristic-curves [14]

The analysis of the speed reduction does not show the complete system performance. Receiver-Operating-Characteristic-curves (ROC-curve) can be used to show more clearly the system behavior, especially the relation between correct and false system interventions for different operating points of an algorithm. The ROC-curves visualizes the relation between the true-positive-rate (1) and the false-positive-rate (2). Fig. 2 shows the needed definitions of the system reactions for the calculation of the true- and false-positive-rate.

$$TP_{Rate} = TP / (FN + TP) \quad (1)$$

$$FP_{Rate} = FP / (TN + FP) \quad (2)$$

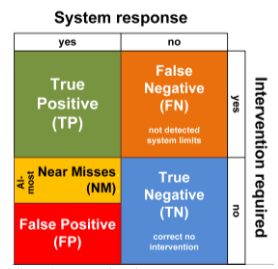

Figure 2. Classification of the system evaluation at systems reactions

- Position of the pedestrian during relevant time steps

The deterministic behavior of the reference algorithm is known because of its implementation, but for the machine learning algorithm it is difficult to identify and validate its behavior. Machine learning algorithms try to identify statistical patterns in the training data during the learning process. One possibility in our case is to analyze the structure of the random forest and the neural network, especially to analyze every tree in the random forest or all nodes in the neural network. But given the high number of nodes in a

TABLE I. OVERVIEW OF SELECTED FEATURES

ID	Features	Time Series
1	Velocity vehicle	x
2	Acceleration vehicle	x
3	x/y-position pedestrian	x
4	Relative x/y-velocity pedestrian	x
5	Velocity pedestrian	x
6	Orientation pedestrian	x
7	Distance vehicle pedestrian	x
8	Angle vehicle pedestrian	x
9	Brake pedal position	-
10-15	Situation meta data	-
16	**Time-to-Collision**[a]	x
17	**Predicted lateral position at TTC=0**[a]	x

[a] Extended features

neural network and the high number of trees and the complexity of every tree this strategy is not feasible. Therefore, we analyze the position of the pedestrian at different time steps to get a better assessment about false system interventions and limitations.

B. Data Basis

Currently no recorded real accident data is available in the required quality. Therefore generated scenarios are being used, which were originally designed for the analysis of the effectiveness of active pedestrian safety system [15]. The generated scenarios include a large number of stochastic simulations, in which a pedestrian is crossing the street. Thereby, critical situations can occur, when the pedestrian estimates the time gap for the road crossing incorrectly, and the driver cannot react in time. About 75% of pedestrian accidents are covered with crossing scenarios from the nearside and far side when the vehicle is traveling straight [16]. Therefore the crossing pedestrian was selected for further considerations. The training data contains 1840 scenarios (242 accidents) and the test data 1829 scenarios (243 accidents).

To investigate the robustness of the machine learning algorithm, one opportunity is to vary the amount of training data. In our case all scenarios where the pedestrian walks between 4 km/h to 6 km/h are removed. Fig. 3 shows the distribution of the amount of pedestrian velocities for the training and test data. The white bars show the not considered scenarios to evaluate the robustness.

IV. Simulation Results

The machine learned algorithms have been compared to a reference system, therefore, the 1829 generated pedestrian scenarios are used. The behavior of the reference system corresponds to the AEB system in series production cars, with some slightly different system parameters.

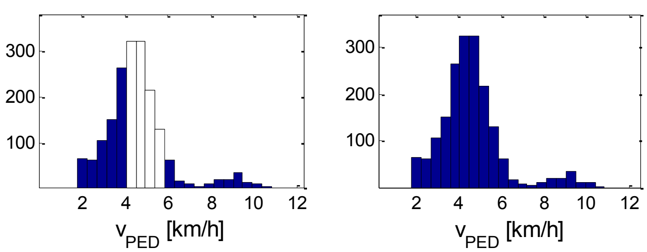

Figure 3. Distribution of the pedestrian velocity for the training data (left) and test data (right)

A. System behavior

Fig. 4 shows the ROC-curves for all systems applied on the 1829 scenarios. The ROC-curves visualize the dependence between the true positive rate and the false positive rate for different operating points (reference implementation: earliest breaking time, machine learning: probability to brake). With the shift of the operating point, it is possible to choose the desired collision speed reduction, and get the resulting true and false positive rate, or vice versa for the considered scenarios.

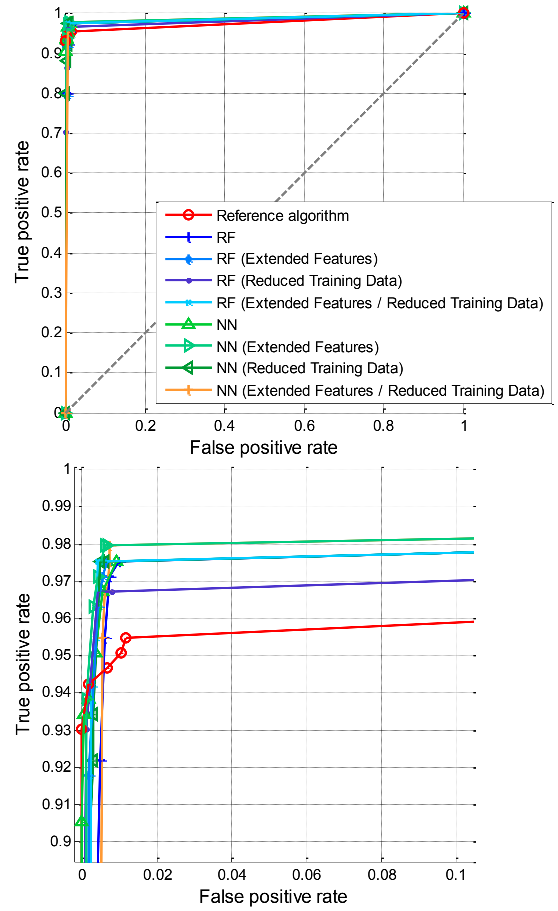

Figure 4. ROC curve for the reference algorithm and the machine learning algorithms. The picture at the bottom is zoomed in to the area of interest.

The ideal operating point would have a true positive rate from 1 and a false positive rate from zero. Fig. 4 clearly shows that all algorithms do not reach this point. This is based on system limitations and the behavior of pedestrians. The vehicle cannot stop immediately and a pedestrian is in some cases able to stop or escape shortly before the collision. All algorithms have a similar behavior. The bottom diagram of Fig. 4 shows, the machine learning algorithms have at the same false positive rate a higher true positive rate and address thereby more critical scenarios.

B. Speed Reduction

To compare both algorithms the ROC-curves by themselves are not sufficient. It does not show the effectiveness of the algorithms. Therefore, an operating point is chosen, in which all algorithms have nearly the same number of false positives. Fig. 5 shows the distribution of the collision speed. The bars at a collision speed of 0 km/h show the amount of avoided accidents for every algorithm. Furthermore, it is interesting that the neural network avoids more collisions and performs better in the complete speed range, without having a higher number of false positives. Detailed results are displayed in Table 2.

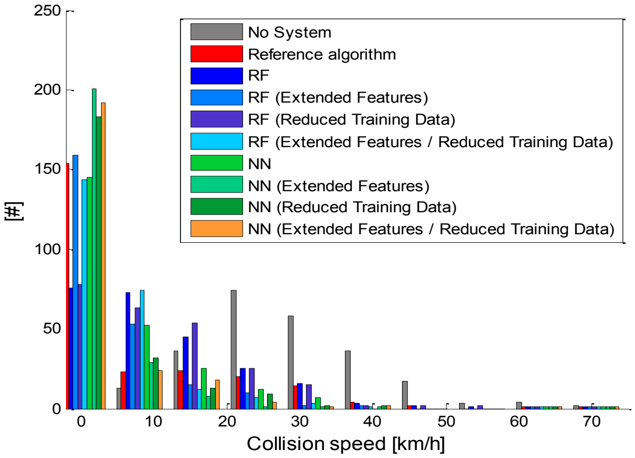

Figure 5. Distribution of the collision speed, with and without active pedestrian protection system.

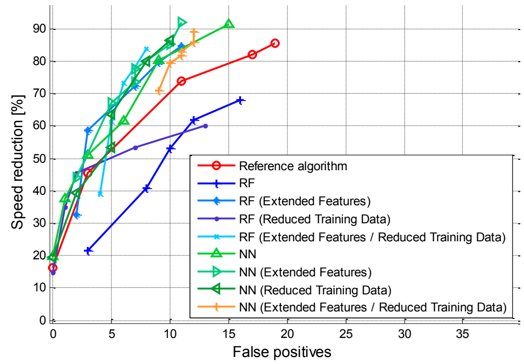

Figure 6. Visualization of the speed reduction in relation to the number of false positives for the reference algorithm and the machine learning algorithms for different operating points.

Another possibility to visualize the effectiveness of the algorithms is the speed reduction in relation to the number of false positives. Fig. 6 shows the results for the reference algorithm compared to different machine learning algorithms. The investigation shows random forests have a lower speed reduction with the standard feature set and the reduced amount of training data compared to the neural networks.

This investigation shows, neural networks can understand the problem better and do not need additional information in the training data, compared to random forests. Also, missing data in between is not a problem. It seems like neural networks in this case learn more the physics of the problem and are therefore able to handle not learned scenarios better. There is no wrong system behavior recognizable.

C. Pedestrian Trajectory

The deterministic behavior of the reference algorithm is known because of the implementation, but for a machine learning algorithm it is difficult to identify and validate its behavior. In our investigation, we evaluate the position of the pedestrian at defined time steps, like the time of the brake request. Fig. 7 shows as an example for true positives and false positives the pedestrian positions at the brake request.

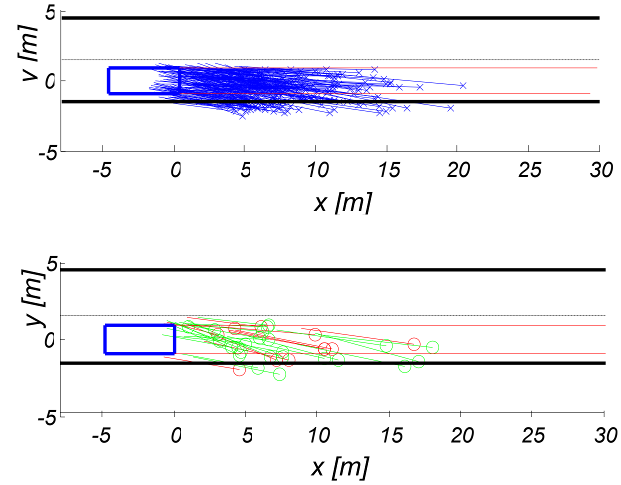

Figure 7. Top: Visualization of the pedestrian positions with the predicted trajectories at the brake request from a Neural Network for true positives. Bottom: Visualization of false positives and near misses at the brake request.

The upper diagram shows the vehicle (blue rectangle), its current path (red line), all positions of the pedestrians (blue crosses) and the calculated predicted trajectories (blue lines) of the pedestrians for true positives at the time of the brake request for the chosen operating point. It can be clearly seen, that all trajectories point towards the vehicle. In some cases the predicted trajectories hit the vehicle front, in these cases the system intervention was later than required to avoid the collision, because the algorithm cannot brake in every situation. This is limited by the system specification.

The bottom diagram shows the calculated predicted trajectories of the pedestrians for false negatives (red) and near misses (green). In the most cases, the linear prediction of the pedestrian trajectory meets the relevant area near the vehicle front in order for the system to respond, but in these

TABLE II. EVALUATION RESULTS: ANALYSIS OF THE EFFECTIVENESS AND SYSTEM REACTION OF THE DIFFERENT ALGORITHMS

Algorithm	*TP*	*FP/ Near miss*	*FN*	*Rel. Speed reduction [%]*	*Avoided accidents*
Reference algorithm	230	11/30	13	73.7	154
Standard Features					
Random Forest	236	12/21	7	61.7	76
Random Forest reduced training data	235	7/16	8	53.4	53
Neural Network	235	9/20	8	80.4	145
Neural Network reduced training data	237	10/28	6	86.5	183
Extended Features					
Random Forest	237	11/21	6	84.7	159
Random Forest reduced training data	237	8/19	6	83.9	144
Neural Network	238	11/24	5	92.0	201
Neural Network reduced training data	238	12/32	7	89.1	192

scenarios the pedestrian was able to avoid the collision, e. g. with a strong movement at the last moment. In our consideration this is counted as a false positive. In the evaluation of these scenarios the most false positives are no “real” false positives, they are more comparable to near misses. This is due to the fact, that in these scenarios the vehicle would just barely miss the pedestrian or the pedestrian would escape at the last possible moment caused by a strong movement. In these critical situations the driver would understand the system reaction.

V. LIMITATIONS

The paper only describes a theoretical analysis of the method, without the consideration of legal aspects. The available scenarios for the investigation only cover a limited action space. If the driver’s evasion maneuver is not included in the action space, the algorithm cannot learn this behavior. It would detect an unknown behavior in a real life situation. In the paper only modeled scenarios for the evaluation are used. A model is only an approximation of the real world and has by definition deviations, an incorrect behavior can be learned from the algorithm. Based on these arguments, real data must be used for a series application, to get a realistic behavior of the learned algorithm.

VI. CONCLUSIONS

The proof-of-concept shown Machine Learning offers the potential to improve Active Safety. It is possible to reach the desired system behavior for the considered traffic scenarios. The machine learning algorithms achieve a higher collision speed reduction and do not trigger more false system reactions. In our investigation, neural networks perform better than random forests and the reference algorithm. Neural networks handle the missing excluded speed range of the pedestrian better and the standard feature set without a performance drop, compared to the random forests. However, the verification of the system behavior places new demands. Active Safety Systems, in particular, must also feature deterministic and comprehensible system behavior. To continue the investigation, we need more accurate traffic data with a higher variation in accident types and causes. Additional generated traffic data can be used to continue investigating our approach for a higher scenario variation. However, a machine-learned Active Safety System for a production car can only be developed based on real-world data. Therefore, we need first Highly Automated Driving vehicles on the road.

VII. ACKNOWLEDGMENT

VIRTUAL VEHICLE Research Center is funded within the COMET - Competence Centers for Excellent Technologies - programme by the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT), the Federal Ministry of Science, Research and Economy (BMWFW), the Austrian Research Promotion Agency (FFG), the province of Styria and the Styrian Business Promotion Agency (SFG). The COMET programme is administrated by FFG. We would furthermore like to express our thanks to our supporting industrial and scientific project partners.

REFERENCES

[1] T. Heinrich, J. Ortlepp, J. Schmiele and H. Voß, "Infrastrukturgestützte Fahrerassistenz," UDV (German Insurers Accident Research), Berlin, 2011.

[2] M. Bojarski, P. Yeres, A. Choromanaska, K. Choromanski, B. Firnder, L. Jackel and U. Muller, "Explaining How a Deep Neural Network Trained with End-to-End Learning Steers a Car," 2017.

[3] F. Netter and F. Friedman, "Artificial Intelligence for Self-Driving Cars: Opportunities & Technologies," Nvidia GTC, Amsterdam, 2016.

[4] M. Huelsen, Knowledge-Based Driver Assistance Systems, Wiesbaden: Springer Fachmedien, 2014.

[5] M. Liebner and F. Klanner, "Driver Intent Inference and Risk Assessment," in *Handbook of Driver Assistance Systems*, Springer, 2015.

[6] M. Goldhammer, S. Köhler, K. Doll and B. Sick, "Camera Based Pedestrian Path Prediction by Means of Polynomial Least-squares Approximation and Multilayer Perceptron Neural Networks," in *SAI Intelligent Systems Conference*, London, 2015.

[7] C. Hyunmin, M. K. Chang, K. ByeoungDo, K. Jaekyum, C. C. Chung and W. C. Jun, "Autonomous Braking System via Deep Reinforcement Learning," in *CoRR*, 2017.

[8] DESTATIS, "Road Accidents 2016," Statistisches Bundesamt, Wiesbaden, 2017.

[9] A. Schubert, C. Erbsmehl and L. Hannawald, "Standardized pre-crash-scenarios in digital format on the basis of the VUFO simulation, Verkehrsunfallforschung an der TU Dresden GmbH," in *Expert Symposium on Accident Research*, Hannover, 2012.

[10] T. Hummel, M. Kühn, J. Bende and A. Lang, An investigation of their potential safety benefi ts based on an analysis of insurance claims in Germany, Berlin: UDV (German Insurers Accident Research), 2011.

[11] Aeberhard N., Rauch S., Bahram M., Tanzmeister G., Thomas J., Pilat Y., Homm F., Huber W., Kaempchen N., "Experience, Results and Lessons Learned from Automated Driving on Germany's Highway," IEEE Intelligent transportation systems magazine, 2015.

[12] M. Schratter, S. Cantu, T. Schaller, P. Wimmer and D. Watzenig, "Technology from Highly Automated Driving to Improve Active Pedestrian Protection Systems," Detroit, 2017.

[13] T. Helmer, T. Kühbeck, C. Gruber and R. Kates, "Development of an Integrated Test Bed and Virtual Laboratory for Safety Performance Prediction in Active Safety Systems," in *SAE-China, FISITA (eds) Proceedings of the FISITA 2012 World Automotive Congress. Lecture Notes in Electrical Engineering, vol 197*, Springer, Berlin, Heidelberg, 2012.

[14] T. Fawcett, "An introduction to ROC analysis," in *Pattern Recognition Letters, Volume 27, Issue 8*, Niederlande, Elsevier, 2006, pp. 861-874.

[15] R. Kates, O. Jung, T. Helmer, A. Ebner, C. Gruber and K. Kompass, Stochastic simulation of critical traffic situations for the evaluation of preventive pedestrian protection systems, VDI Verlag GmbH, 2010.

[16] C. Domsch, Leistungssteigerung präventiver Schutzsysteme, München: 5. Tagung Fahrerassistenz, Fahrzeugsicherheit, 2012.