Hannes Weinreuter*, Jonas Imbsweiler, Nadine-Rebecca Strelau, Barbara Deml, and Fernando Puente León

Prediction of human driver intentions at a narrow passage in inner city traffic

Intentionsprädiktion menschlicher Fahrer an einer Engstelle im innerstädtischen Straßenverkehr

DOI 10.1515/teme-2019-0063

Abstract: Autonomous vehicles have to be able to predict whether a human driver will wait at an unregulated inner city narrow passage or not to adapt its behaviour accordingly. To this end, a driving simulator study was conducted in which participiants were subjected to different cooperation behaviours during their approach to a narrow passage. They were asked to rate their intention afterwards. From the recorded trajectories, features which are specific to the scenario are derived. Therewith, *Random Forest* and *Conditional Random Field* classifiers for both intention and behaviour prediction are trained. The results show that robust prediction of driver intention and behaviour is possible.

Keywords: Autonomous driving, intention prediction, machine learning.

Zusammenfassung: Autonome Fahrzeuge müssen vorhersagen können, ob ein menschlicher Fahrer an einer innerstädtischen, nicht regulierten Engstelle warten wird oder nicht, um ihr Verhalten entsprechend anzupassen. Zu diesem Zweck wurde eine Fahrsimulatorstudie durchgeführt, in welcher die Probanden verschiedenem Kooperationsverhalten während ihrer Anfahrt auf eine Engstelle ausgesetzt wurden. Anschließend wurden sie gebeten, ihre Intention einzustufen. Anhand der aufgezeichneten Trajektorien wurden szenariospezifische Merkmale abgeleitet. Mit diesen wurden *Random-Forest-* und *Conditional-Random-Field*-Klassifikatoren sowohl für die Intention als auch für das Verhalten trainiert. Die Ergebnisse zeigen, dass die robuste Vorhersage der Intention und des Verhaltens eines Fahrers möglich sind.

Schlüsselwörter: Autonomes Fahren, Intentionsprädiktion, maschinelles Lernen.

1 Introduction

In order to safely integrate automated vehicles into traffic, several challenging tasks are yet to be resolved. One of these challenges is inner city traffic. In addition to the fact that there are often many traffic participants with conflicting interests, there are some situations that are not clearly regulated by law. These situations include symmetrical narrow passages (see Fig. 1) and unregulated junctions (*right before left* applies). In these situations, no one has the right of way and drivers therefore need to communicate with each other in order to resolve the situation. If a solution is found while the vehicles involved are still approaching the situation, it is possible to avoid a blockage and traffic can flow more efficiently. Situations like these are especially difficult for automated vehicles. They have to be able to interpret human communication signals, decide how to behave and communicate that decision in time to optimize the overall traffic flow.

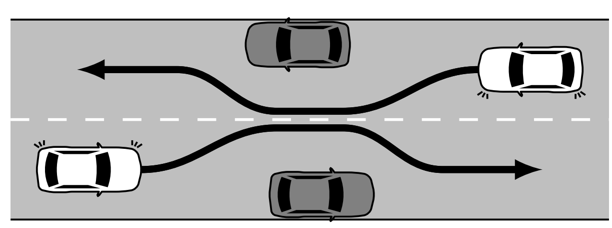

Fig. 1: A symmetrical narrow passage.

Phillips et al. [9] use a *Long Short-Term Memory Network* to predict whether a driver will turn right or left or go straight at an upcoming intersection. Koide and Miura

***Corresponding author: Hannes Weinreuter,** Institute of Industrial Information Technology, Karlsruhe Institute of Technology, Hertzstr. 16, 76187 Karlsruhe, e-mail: hannes.weinreuter@kit.edu
Jonas Imbsweiler, Nadine-Rebecca Strelau, Barbara Deml, Institute of Human and Industrial Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 4, 76131 Karlsruhe, e-mail: jonas.imbsweiler@kit.edu, nadine-rebecca.strelau@kit.edu, barbara.deml@kit.edu
Fernando Puente León, Institute of Industrial Information Technology, Karlsruhe Institute of Technology, Hertzstr. 16, 76187 Karlsruhe, e-mail: puente@kit.edu

[6] use *Hidden Conditional Random Fields* to predict if a pedestrian is aware of an obstacle. Their classifier is based on the walking trajectory. Hu et al. [2] propose a *Semantic-based Intention and Motion Prediction* that adapts to different scenarios to predict the final location and the associated timing of surrounding vehicles. Hubmann et al. [3] use a *Partially Observable Markov Decision Process* to output the optimal acceleration assuming the intention of the other vehicles as hidden states. Their model predicts the turning direction of the other vehicles during the decision making as well. Tran et al. [10] predict several manoeuvrers including stopping, lane changes and turning with an *Hidden Markov Model*. They use the vehicle trajectory as well as the state of the steering wheel and the pedals. Lee et al. [8] predict lane change manoeuvrers with a *Convolutional Neural Network*. The prediction is used for motion control of an adaptive cruse control system.

This work focuses on the intention prediction of human drivers at a narrow passage. Here intention is defined as the will to *go first* or to *not go first*. The classifiers are trained with a data set that was generated from a driving simulator study. This approach allows to generate a large data set in a controlled environment compared to observations of real traffic. All participants drove through the same scenarios, which featured different driving styles of the other vehicle. Two types of labels are used: The behaviour, e.g. which driver drove first, and the intention of the drivers, e.g. what did the participant want to do while approaching the narrow passage.

The paper is structured as follows: In Section 2 the data set that has been used for this work is introduced. In Section 3 the proposed intention prediction algorithm is described. In Section 4 the prediction performances of the classifiers when applied to the data set are presented. Section 5 concludes the paper and gives a summary and an outlook to future work.

2 Data set

The data set for this work originates from a driving simulator study. The simulator itself consists of the front half of a passenger vehicle that includes the driver's and passenger's seat. The projection screen is curved and extends to the side windows. It therefore covers the entire field of view of a person sitting in the driver's seat.

In the study participants drove through six narrow passage and five T-junction scenarios. The cooperation vehicles were set up to arrive at the scenario at the same time as the participant. These vehicles (one for the narrow passage and two in the case of the T-junction) were not present on the course at all times, instead they were set to appear and disappear out of sight of the participants. The vehicles appeared at a set distance, the appearance was triggered by the participant driving over a trigger point at the same distance from the scenario. After that their velocities were synchronised with the participant's velocity at regular intervals, thus ensuring a simultaneous approach.

Fig. 2: Narrow passage in the driving simulator.

After the last synchronisation the vehicle's behaviour depended on the current script and was either offensive or defensive. The cooperation vehicles followed their specified behaviour regardless of the participant's behaviour. Only if the participants entered the scenario at the same time as the cooperation vehicles, an emergency stop of these vehicles was triggered. The six narrow passage scripts are described in Table 1. Similar scripts were used for the T-junction. Both the narrow passage and the T-junction scripts were adapted from an earlier study by Imbsweiler et al. [4]. Every participant drove through all of the eleven scripts once in a randomised order. Figure 2 shows a screen shot of the interaction at the narrow passage, created by the two stationary yellow vehicles, in the simulator.

Table 1: Narrow passage scripts.

Script	Description	Category
1	stop distinctively	defensive
2	reduce velocity & flash headlights	defensive
3	stop distinctively & flash headlights	defensive
4	unchanged velocity	offensive
5	accelerate	offensive
6	decelerate	offensive

After each scenario, they were asked to stop and answer a questionnaire about their intention during the approach to the scenario. In total 29 people successfully participated in the study. In this work only the narrow passage data is used. The trajectories were recorded with a frame rate of 60 Hz and serve as input data for training and validation of the classifiers for intention prediction.

3 Intention prediction

In this work the intention is predicted using the driven trajectories $\mathbf{r_o}(t) = [x_\mathrm{o}(t), y_\mathrm{o}(t)]^\mathrm{T}$ of the observed vehicle and the trajectories of the ego vehicle $\mathbf{r_e}(t) = [x_\mathrm{e}(t), y_\mathrm{e}(t)]^\mathrm{T}$. In the context of this paper the observed trajectory always refers to the participant's trajectory since only the intention of the participant is of interest. From these trajectories a feature set is derived. The features were chosen to include both generic parameters of driving and features that are specific to a narrow passage scenario. The following features are used in this work:

The *longitudinal acceleration* $a_\mathrm{lon,o}$ is the acceleration of the observed vehicle along the current vehicle heading. The averaged feature is shown in Figure 3. The features are plotted separately for the scripts with the vehicle showing offensive and defensive behaviour.

The *distance to scenario* $d_\mathrm{s,o}$ is the distance from the current position of the observed vehicle to the front of the narrow passage along the centre of the lane.

The *weighted lateral deviation* $d_\mathrm{lat,o}$ describes the distance from the lane centre to the current vehicle position. This distance is measured along the normal to the lane centre through the current position. This feature is normalized such that a vehicle whose centre is on edge of its lane has a lateral deviation of $d_\mathrm{lat} = \pm 1$. Additionally, this feature is weighted with the normalised distance to the narrow passage $d_\mathrm{s} = 1 - \frac{d_\mathrm{s,m}}{d_\mathrm{e}}$ and a scaled logistic function [5]:

$$d_\mathrm{lat} = \frac{d_\mathrm{lat,m}}{0.5 \cdot w_\mathrm{l} d_\mathrm{s}} \frac{1}{1 + \exp\left(-k_1 \left(d_\mathrm{s} - k_2\right)\right)} \,. \qquad (1)$$

$d_\mathrm{lat,m}$ is the measured lateral deviation, $d_\mathrm{s,m}$ is the measured distance to the narrow passage, w_l is the lane width, d_e is the evaluation length for this feature, k_1 and k_2 are the parameters to scale the logistic function. The lateral deviation is weighted to suppress effects from a curve that was placed just before one side of the narrow passage in our simulator set up. This curve was cut by many participants which led the classifiers to recognise the scripts rather than the behaviour or the intention. This effect is amplified by the fact that all offensive scripts were run

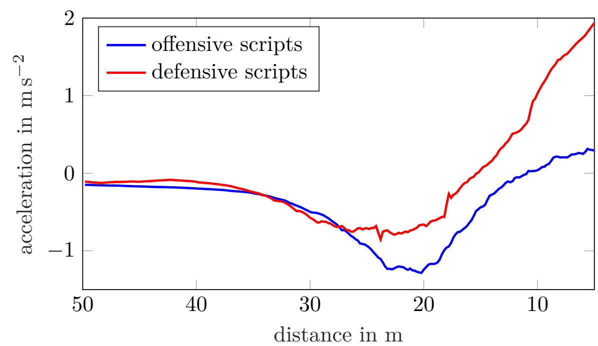

Fig. 3: Averaged longitudinal acceleration feature of the participants by defensive (scripts 1 to 3) and offensive (scripts 4 to 6) behaviour of the simulated vehicle.

such that the simulated car came from the same side. The weighting parameters were chosen such that the feature does not differ for offensive and defensive scripts until a distance of approximately 25 m. This is similar to features that do not show that phenomenon like the longitudinal acceleration a_lon (see Fig. 3). At this distance from the narrow passage the road is straight again.

The *velocity quotient* v_q is the absolute velocity of the observed vehicle $v_\mathrm{abs,o}$ relative to the absolute velocity of the ego vehicle $v_\mathrm{abs,e}$:

$$v_\mathrm{q} = \frac{v_\mathrm{abs,o}}{v_\mathrm{abs,e}} \,. \qquad (2)$$

The features are calculated for each time step, the feature set $\mathbf{F}$ therefore consists of a set of 4 individual feature vectors $\mathbf{f}_k$ of length T:

$$\mathbf{F}(t) = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\} = \{\mathbf{a}_\mathrm{lon,o}, \mathbf{d}_\mathrm{s,o}, \mathbf{d}_\mathrm{lat,o}, \mathbf{v}_\mathrm{q}\}, \qquad (3)$$

$$\mathbf{f}_k = [f_k[1], \ldots, f_k[T]]^\mathrm{T} \,. \qquad (4)$$

These features serve as input to the classification algorithms. In this work, two types of algorithms are used: direct and sequential classifiers.

Direct classifiers require vectors of scalar features as input. To convert the feature set $\mathbf{F}$ into scalar features, it is split into segments of 0.25 s length. To generate the vector of scalar features $\mathbf{f^s}$, the features of a segment are averaged over time:

$$\mathbf{f^s} = \{f_1^\mathrm{s}, f_2^\mathrm{s}, f_3^\mathrm{s}, f_4^\mathrm{s}\} = \{a_\mathrm{lon,o}^\mathrm{s}, d_\mathrm{s,o}^\mathrm{s}, d_\mathrm{lat,o}^\mathrm{s}, v_\mathrm{q}^\mathrm{s}\} \,. \qquad (5)$$

With f_k^s being the arithmetic mean over a segment of the k-th feature. The models are then trained and evaluated on the individual segments.

There are several classifiers that fall into that category; in this work *Random Forests (RF)* [1] are used. A *RF* utilizes several decision trees that are trained with a subset of the data set. Additionally, each branch in a tree only considers a subset of all available features [1].

Sequential classifiers directly use the feature vectors $\mathbf{f}_k$ of the feature set $\mathbf{F}$ as input and output a prediction for every time step. *Conditional Random Fields (CRF)* [7] are an example of such a classifier and are used in this work.

4 Results

From the trajectories obtained by the study, the features are calculated and used to validate the proposed intention prediction system. To this end, separate classifiers were trained for both intention and behaviour.

During the study, participants were asked to rate their intention during the approach to the scenario on a seven-valued scale from *absolutely want to go first* to *absolutely do not want to go first*. Due to the small size of the data set, the intention was sampled down to two answers, *want to go first* and *do not want to go first*. Runs with a neutral label were omitted from the data set, the remaining runs were assigned to the closer one of the two labels. In total, 172 runs remained with 71 runs labeled as *want to go first*.

The behaviour label was obtained from observation and all runs fell into one of the categories *went first* and *did not go first*. Here, a total of 188 runs could be used for classification of which 80 were labeled as *went first*.

Because any prediction has to be performed before the vehicle enters the narrow passage, the features have to be cut so that they only include data from the approach prior to entering. It is also reasonable to assume that a human driver only makes a decision when the vehicle is close enough to the scenario. For these reasons, the features are cut using different start distances $d_\mathrm{s} = \{20\,\mathrm{m}, 25\,\mathrm{m}, 30\,\mathrm{m}\}$ and end distances $d_\mathrm{e} = \{5\,\mathrm{m}, 7.5\,\mathrm{m}, 9\,\mathrm{m}\}$.

Using three-fold cross validation, *CRF* and *RF* classifiers are trained for both intention and prediction. The classifiers were optimized for best performance at the last segment cut out of the feature sequence (for *RF*) or at the last label (for *CRF*).

The performance is evaluated with the f-measure [5]:

$$f = 2\frac{pr}{p+r}\,, \tag{6}$$

$$p = \frac{tp}{tp+fp}\,, \quad r = \frac{tp}{tp+fn}\,. \tag{7}$$

With p being the precision, r the recall and tp the true positive, fp the false positive and fn the false negative values.

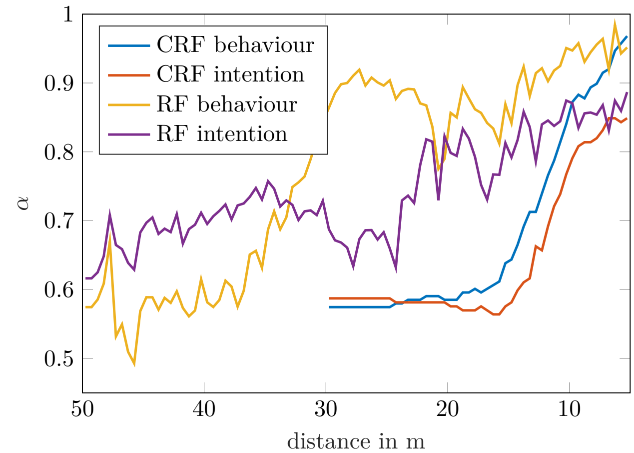

Fig. 4: Classifier performance over distance to narrow passage.

The performance of the classifiers for all training intervals is shown in Table 2. Additionally, the archived accuracy α is also given for comparison. Both the intention and the behaviour can reliably be predicted using both algorithms. The results in general are better if the training interval ends close to the narrow passage. It is evident that the behaviour can be predicted more reliably than the intention for this data set. The results finally indicate that *RF* performs better than *CRF*, reaching a $f = 0.984$ compared to $f = 0.973$ for the best behaviour prediction.

To evaluate the performance of the classifiers over the distance, the best classifiers of each type were further analysed. The classifiers trained with the features from 30 m to 5 m were used in all cases but for the intention prediction with *RF*, in the latter case the classifier trained with features in the range of 20 m to 5 m was used. For the evaluation, the distance from 50 m to 5 m was split into bins with a length of 0.5 m.

Using the selected *RF* classifier, the predicted feature segments were then assigned to the bins to obtain the performance over the distance. The performance of the *CRF* was evaluated by applying the trained model to a feature interval starting at the current bin center. The interval has the same length as the best learning interval (in this case 25 m). As before, the last predicted label of the sequence determined the prediction. The results of that analysis are shown in Fig. 4.

All classifiers show an improved performance while approaching the narrow passage. The *RF* classifiers show superior performance than the *CRF* classifiers, however, since the segments are evaluated individually, their performance does not increase steadily as do the *CRF* versions. At the smallest distance to the narrow passage, the per-

Table 2: Prediction accuracy for intention and behaviour with the best performing intervals marked in green and the worst in orange.

Start distance d_s	End distance d_e	Intention *RF*		Intention *CRF*		Behavior *RF*		Behavior *CRF*	
		f	α	f	α	f	α	f	α
20 m	5 m	0.917	0.919	0.874	0.878	0.984	0.984	0.973	0.973
25 m	5 m	0.904	0.907	0.874	0.878	0.978	0.979	0.973	0.973
30 m	5 m	0.894	0.895	0.880	0.884	0.984	0.984	0.973	0.973
20 m	7.5 m	0.894	0.895	0.855	0.860	0.946	0.947	0.928	0.931
25 m	7.5 m	0.859	0.860	0.830	0.837	0.952	0.952	0.917	0.920
30 m	7.5 m	0.876	0.878	0.842	0.849	0.946	0.947	0.912	0.915
20 m	9 m	0.864	0.866	0.836	0.843	0.952	0.952	0.913	0.915
25 m	9 m	0.881	0.884	0.840	0.849	0.941	0.941	0.890	0.894
30 m	9 m	0.876	0.878	0.829	0.837	0.946	0.947	0.895	0.899

formance of both *CRF* and *RF* are very similar with the classifiers for the behaviour outperforming the intention classifiers. The results are not fully identical with those shown in Table 2 due to the binning.

5 Summary & outlook

In conclusion, it can be said that it is possible to reliably predict the behaviour and the intention of a human driver at a narrow passage. Both *CRF* and *RF* show similar results close to the narrow passage; at greater distances *RF* provide more reliable predictions. In all cases the behaviour is predicted more robustly than the intention. Several conclusions for including this approach in a decision making algorithm for an autonomous vehicle at a narrow passage can be drawn from these results. In order to get reliable classifiers, features with small distances to the obstacles have to be included in the training interval. Especially with *RF*, predictions at greater distances are possible allowing multiple predictions during the approach.

In future work, a greater data set containing different narrow passages could be used to further generalize the classifiers. More features could also be included to make the prediction more robust. The approach should also be extended to include more situations that require intention prediction for autonomous driving, such as unregulated intersections. Finally, a decision making algorithm can be equipped with these classifiers for real time tests in a driving simulator and eventually in a real vehicle.

Acknowledgment: This project has been funded within the priority program 1835 "Cooperatively Interacting Automobiles" by the *German Reasearch Foundation (DFG)*. Mapping data has been provided by the city of Karlsruhe. The authors thank the project partners for their fruitful cooperation.

References

[1] J. Beyerer, M. Richter, and M. Nagel. *Pattern Recognition: Introduction, Features, Classifiers and Principles*. Walter de Gruyter, Berlin, 2017.

[2] Y. Hu, W. Zhan, and M. Tomizuka. Probabilistic prediction of vehicle semantic intention and motion. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 307–313.

[3] C. Hubmann, M. Becker, D. Althoff, D. Lenz, and C. Stiller. Decision making for autonomous driving considering interaction and uncertain prediction of surrounding vehicles. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pages 1671–1678.

[4] J. Imbsweiler, R. Palyafári, F. Puente León, and B. Deml. Untersuchung des Entscheidungsverhaltens in kooperativen Verkehrssituationen am Beispiel einer Engstelle. *at – Automatisierungstechnik*, 65(7):477–488, 2017.

[5] S. W. Knox. *Machine learning: a concise introduction*, volume 285. John Wiley & Sons, 2018.

[6] K. Koide and J. Miura. Estimating person's awareness of an obstacle using HCRF for an attendant robot. In *Proceedings of the Fourth International Conference on Human Agent Interaction*, pages 393–397. ACM, 2016.

[7] J. D. Lafferty, A. McCallum, and F. C. N. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, pages 282–289, San Francisco, CA, USA, 2001.

[8] D. Lee, Y. P. Kwon, S. McMains, and J. K. Hedrick. Convolution neural network-based lane change intention prediction of surrounding vehicles for ACC. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–6.

[9] D. J. Phillips, T. A. Wheeler, and M. J. Kochenderfer. Generalizable intention prediction of human drivers at intersections. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pages 1665–1670.

[10] D. Tran, W. Sheng, L. Liu, and M. Liu. A hidden markov model based driver intention prediction system. In *2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, pages 115–120.