The PREVENTION Challenge: How Good Are Humans Predicting Lane Changes?

A. Quintanar, R. Izquierdo, I. Parra, D. Fernández-Llorca, and M. A. Sotelo

Abstract— While driving on highways, every driver tries to be aware of the behavior of surrounding vehicles, including possible emergency braking, evasive maneuvers trying to avoid obstacles, unexpected lane changes, or other emergencies that could lead to an accident. In this paper, human's ability to predict lane changes in highway scenarios is analyzed through the use of video sequences extracted from the PREVENTION dataset, a database focused on the development of research on vehicle intention and trajectory prediction. Thus, users had to indicate the moment at which they considered that a lane change maneuver was taking place in a target vehicle, subsequently indicating its direction: left or right. The results retrieved have been carefully analyzed and compared to ground truth labels, evaluating statistical models to understand whether humans can actually predict. The study has revealed that most participants are unable to anticipate lane-change maneuvers, detecting them after they have started. These results might serve as a baseline for AI's prediction ability evaluation, grading if those systems can outperform human skills by analyzing hidden cues that seem unnoticed, improving the detection time, and even anticipating maneuvers in some cases.

Fig. 1. PREVENTION Challenge overview.

I. INTRODUCTION AND RELATED WORK

According to [1], about 8% of road fatalities in the EU occurred on highways, which represents more than 2k people involved. This number must be reduced in upcoming years to meet the proposed objectives [2], trying to halve the number of serious injuries. Lane change accidents represent one of the leading causes of motorway accidents. A lane change accident occurs when a vehicle maneuvers laterally from one lane into another, colliding with a vehicle that was already in the destination lane earlier. Nowadays, many manufacturers develop driving aid systems that allow, for example, to determine whether a vehicle is in the blind spot or not, avoiding a possible emergency. Due to the forced coexistence of vehicles equipped with these systems and those lacking them, there is a need to conduct a study that analyses whether humans are capable of detecting and avoiding a lane change accident in advance.

The use of turn signal is another primary cue used to infer whether a driver desires to start a lane-change maneuver. As stated in [3], the use of turn signals may vary significantly among drivers, and evaluating data on turn signal could lead to improve prediction schemes. In some cases, drivers that use turn signals activate them after beginning the lane-change maneuver, as shown in [4]. Our study carefully assesses the lane-change data contained in The PREVENTION Dataset [5], which includes examples of most common types of

A. Quintanar, R. Izquierdo, I. Parra, D. Fernández-Llorca and M. A. Sotelo are with the Computer Engineering Department, Universidad de Alcalá, Alcalá de Henares, Spain. alvaro.quintanar@uah.es

maneuvers, including when a driver changes lanes to pass a slower lead vehicle to maintain current speed, a typical situation in Spanish highways.

To detect lane changes, drivers should identify areas of risk approaching and focus on them, e.g., a merge action in a busy highway. Watching for subtle changes in other vehicles' motion to identify the drivers' intent, such as drifting in the lane or slight angling of the front wheels, could help drivers avoiding critical situations. A vehicle driver will predict the intentions of other traffic participants to decide to anticipate and evade a potentially dangerous conflict. Thus, correct predictions ease a quick reaction, but incorrect ones could be potentially hazardous.

According to [6], a lane-change maneuver is desirable if it is either discretionary, anticipatory, or mandatory (initialized by road conditions). The sequences included in the study display different lane-change maneuvers, including double-triple simultaneous LCs, hazardous maneuvers, emergency vehicles in traffic jams and other circumstances.

A. Situational awareness and scene understanding

Situational awareness (SA) could be defined as the ability to scan the environment and perceive danger, challenges, and opportunities while maintaining the ability to conduct normal activities. Driving in traffic consists of recognizing and reacting to potentially dangerous scenarios, maneuvering efficiently through traffic, and achieving higher-level objectives while addressing other concerns [7].

By receiving advanced driver training [8], drivers can develop and improve their SA, but not necessarily in the way that is intended. Thus, it is clear that a holistic method to identify and understand every cue present in a specific situation can help drivers to react in those extreme conditions where an accident is near to happen.

In [9], lane changes are classified into five categories regarding the interaction style observed with surrounding vehicles. Most lane-change maneuvers analyzed in this study were self-motivated, and no significant interaction with other vehicles was observed. SA is necessary to understand and react in complex situations, as several vehicles changing lanes at the same time or a congested highway with a vehicle fast approaching that wants to overtake.

Regarding driving skills of novice drivers, in [10] young drivers seemed to recognize that they were not as skilled as the average driver . Moreover, overconfident drivers were worse at adapting their behavior to the traffic situation, acting more aggressively. Comparing young and old drivers' behavior in [11] result in the need for old drivers to check their surroundings before performing a lane change. In a dynamic driving, the environment is critical to detect and perceive cues and analyzing the future state of the driving agents based on those cues to react appropriately and anticipate hazardous maneuvers. For old drivers, the post-drive feedback was useful in terms of increasing their situational awareness.

B. Vehicle intention prediction

Apart from analyzing the scene by itself, in this study, the *holistic image* of every user is assessed by comparing the driving situation understanding capabilities, predicting other vehicles' lane changes in a great set of sequences. As in [12], overtake-late (lane crossing) and overtake-early (initial lateral motion) maneuvers are compared with the very moment when the user considers that a lane-change maneuver is starting.

In this paper, the PREVENTION dataset [5] has been used to assess human capabilities in detecting and predicting lane changes in highway scenarios. The dataset offers more than 356 minutes, 4 million vehicle detections, and 3000 trajectories, collected from up to 8 sensors, including LiDAR, RADAR, DGNSS, and cameras, acquiring information from surrounding vehicles up to a range of 100 meters. Among other data, frontal camera information along with both manual and CNN detections, as well as lane-change annotations have been used to create the Challenge. It is essential to notice that participants assessed lane-change maneuvers of surrounding vehicles, not of the ego vehicle. The main objective of this work is to prove if participants are able to anticipate lane-change maneuvers or not.

The paper is organized as follows: Section II offers an exhaustive description of the methodology applied to create the PREVENTION Challenge, by selecting and adapting the sequences to perform the trials, as well as developing a questionnaire to assess users' demographic information regarding the survey. In Section III, several results are described,

creating a regression model and analyzing if humans are able to predict lane changes in highway scenarios. Section IV offers a discussion about the results and the Challenge itself, by evaluating the content, trials, and the contribution to the state of the art. Finally, Section V completes the paper, providing conclusions extracted from the Challenge and suggesting some future developments of prediction of vehicle intention.

II. METHODOLOGY

This section describes the methodology employed to perform the survey, starting by selecting the participants, designing the architecture, revising and fine-tuning the materials used and the process of trial, post-processing the data, and finishing by extracting some possible predictive factors and statistics about human capabilities compared with previously labeled data.

A. Participants

Participants were recruited among the Engineering School, from students to teachers and other research staff, as well as family, friends, and colleagues. Thus, more significant variability is achieved in terms of the expertise of the participants.

B. Architecture

The study was performed using a Qt-based application, where 30 randomly selected sequences were presented to the participants to select the starting instant of a lane change or do nothing in its absence. The vehicle of interest was inside a red bounding box; participants were asked to hit any key as soon as they notice an upcoming lane change maneuver performed by that vehicle. The independent variable was defined as the lane change existence and its direction (three possible options: left lane change, right lane change, or none). If the participant did not spot any lane change, the sequence would continue up to its finishing point, and that answer was marked as a "none". The dependent variables were defined as prediction correctness and prediction certainty. Some insights about the application can be observed in Figure 2.

C. Materials: revision and fine-tuning

A total of 794 different sequences were selected from the PREVENTION dataset. The dataset offers different sequences recorded in various situations and locations, with precise labeling of lane changes. The sequences are video fragments taken from the front camera, recorded at ∼10 Hz, and conveniently cropped to 1920x600, removing the top and bottom bands that contain irrelevant information. Recordings were made in a naturalistic way, driving in real traffic during central hours, avoiding heavy road traffic. As the lane changes are already labeled, including blinker use during the maneuver, all lane-change maneuvers were carefully revised, and those that seemed to be unclear for humans to understand, such as far vehicles or ambiguous IDs, were conveniently removed from the trial set. The total amount of LC maneuvers used in the trials and their

Fig. 2.   Application outline.

nature are described in Table I. Nevertheless, the survey also needed scenes where there were vehicles that keep their trajectory in the lane, that is, sequences without lane-change maneuvers of a specific vehicle. To do this task, the authors revised every recording, labeling appropriate fragments of different vehicles, obtaining the sequences of no-lane-change maneuvers described in Table I.

TABLE I

MANEUVERS COUNT

[a] NO LANE CHANGE, [b] LEFT LANE CHANGE, [c] RIGHT LANE CHANGE

Record # Drive #	NLC[a]	LLC[b]		RLC[c]	
		$\overline{blinker}$	blinker	$\overline{blinker}$	blinker
1 - 1	11	4	15	7	38
2 - 1	7	1	11	4	15
2 - 2	13	0	13	2	18
3 - 1	14	2	12	3	11
3 - 2	14	2	7	0	19
4 - 1	29	5	23	17	59
4 - 2	24	9	39	15	42
4 - 3	7	1	7	1	9
5 - 1	20	6	24	12	46
5 - 2	20	3	35	9	39
5 - 3	20	1	17	1	11
Total	179	34	203	71	307
		237		378	
%	22.54%	4.28%	25.57%	8.94%	38.66%
		29.85%		47.61%	

As there were up to 800 sequences in total, it was necessary to select a few of them for each trial randomly. Thus, all lane change annotations were stacked, filtered, and shuffled randomly, selecting 10 of each type to form the questions for a single participant. To avoid showing the same video of a sequence (to increase variability), a random number of previous frames (between 50 and 100) are showed. Balance of maneuvers type and their selection is discussed in Section IV-A; in Figure 3 various sequences examples are shown.

Regarding making the participant focus on the selected vehicle, firstly, high-level labeling was developed, including the reference of the lane with respect to the ego and the color of the vehicle of interest. After a few test volunteers, that idea was deprecated, and another method based on the bounding box was selected: the vehicle of interest is marked with a red bounding box with a thickness of one pixel, leaving the scene almost unaltered to identify a possible cue that would make the user infer a lane-change maneuver.

Fig. 3.   Main frames of various sequences.

D. Collection of demographic information

Participants were asked to provide demographic details, such as age, gender, occupation, and whether they have a driving license. If so, they were asked additionally about their experience (less than a year, between one and two years or more than two years), their driving frequency (daily, only during weekends or occasionally) and their typical driving zones (urban scenarios or mostly highways and two-way roads). Moreover, optionally they could input their full name and e-mail address if they wished to be notified about the study results. The question about driving license is crucial to distinguish between actual drivers that could recognize typical traffic situations and those who not, including graduation regarding their experience, driving frequency, and driving scenarios. For example, frequent and experienced drivers who are familiar with highway scenarios will likely predict a merging maneuver before a participant who does not have a driving license.

E. Trials

A fixed procedure was developed to proceed with every volunteer:

1) After filling up the demographics form, together with the user, a single ID is issued, which is going to be

introduced in the Challenge application, as it is shown in figure 2.

2) After clicking "Start Test", a prompt window is showed, explaining the process to the user. The participant is asked to tell the staff what is needed to do, as well as if there is any question.

3) The prompt window previous to every sequence appears, waiting for the user to click "Accept" in order to start the first sequence.

4) The sequence is displayed, with the vehicle of interest being focused by a red bounding box. If the user detects an upcoming lane-change maneuver, any key will be hit (the volunteer is encouraged to use the space bar).

5) The sequence stops, and the lane change selection screen is showed, formed by two buttons: "Left Lane Change" and "Right Lane Change". The user clicks on the selected option.

6) The prompt window previous to every sequence is showed, asking the user to accept to start the next sequence.

7) When all sequences have been finished, a results window is showed, featuring:
   - Accuracy (%).
   - Delay in detecting the lane change (s).
   - Lane change anticipation (s).

8) The start window is showed, waiting for a new user ID to start a test.

The Challenge is supposed to be done in a controlled environment, so the staff used laptops with the same screen resolution and controllers for every test. They were instructed to follow the process meticulously, making sure that the participant has understood every step and the purpose of the test. In order to avoid unnecessary waiting times, a parallel thread is used to load the frames of the sequence, which leads to smoother user experience.

1) Labeling method: Lane-change maneuvers are labeled in the dataset as follows:

- f0: the frame where the driver starts the maneuver, either activating the turn signal or by showing changes in the trajectory of the vehicle. These labels have been double-checked to assure that they are correct, representing the initial point of the lane change.
- f1: the frame where the rear middle part of the vehicle is just between the lanes; at this point, the lane-change maneuver is virtually ended from the drivers' perception.
- f2: the frame where the vehicle is positioned in the new lane, the lane-change maneuver is completed.

F. Data analysis

The result of the test of every participant is coded in a text file, in order to make post-processing and analysis easier. Among others, ground truth annotations for every lane-change step, type of maneuver, and test results (frame selected by the user and kind of maneuver) are included.

Thus, the post-processing stage consisted of extracting both data from the test application and demographics form, performing the statistical model creation, including histogram and curve fitting with MATLAB.

III. RESULTS

The Challenge was conducted over several weeks, starting in January 2020. A total of 72 participants were tested, assessing 30 randomly selected sequences of highway scenarios. This section presents different results regarding prediction success, regression model, cue analysis, and individual factors as to anticipation of lane-change maneuvers.

A. Prediction success

The dataset used was recorded around Madrid, in the center of Spain, so, logically, local users that maybe commute daily to get to their work or studies could be familiar with them. Some of them stated that during the trial process, being curious about the recording areas. After processing the results, we find it appropriate to remove the worst and the best trial of each participant to avoid biasing or random variation. The measures of delay and anticipation only include hits (that is, when the user response is the same as the ground truth), so the accuracy and missed maneuvers are analyzed in this section as well. Anticipation and delay are converted from frames to seconds, taking into account that the sequences are recorded at ∼10 Hz.
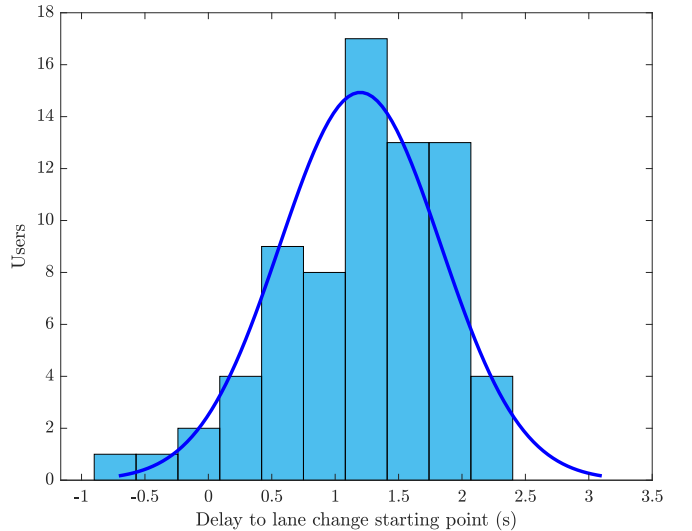

Fig. 4.   Delay to lane-change maneuver starting point.

1) Delay to lane-change starting point: The delay of users to lane change starting point has been defined as the difference between the frame where the user decided that the maneuver was starting and the actual frame f0. A histogram of users regarding the delay is observed in Figure 4, including the fit of a normal distribution ($\mu = 1.198, \sigma = 0.634$) As regard, the mean is above 1.19 seconds; the majority of users detect lane changes after they have already started when the vehicle is clearly changing its trajectory.

2) Anticipation to lane change middle point: The anticipation of users to lane change middle point has been defined as the difference between the frame where the user decided

that the maneuver was starting and the actual frame `f1`. A histogram of users regarding the anticipation is observed in Figure 5, including the fit of a normal distribution ($\mu = 1.680, \sigma = 0.632$). As displayed, the users could at least detect the lane-change maneuver within a reasonable time.
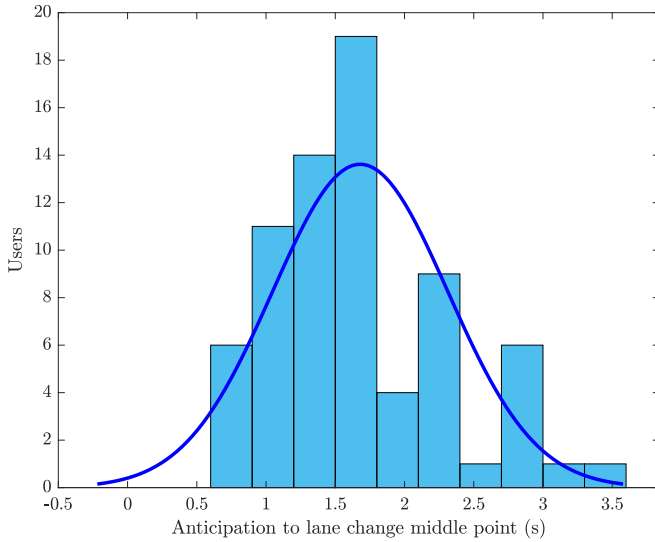

Fig. 5. Anticipation to lane-change maneuver middle point.

3) Delay-accuracy ratio: To extract a conclusion it is necessary to analyze the ratio between the delay to lane change starting point and accuracy. Intuitively, human beings tend to challenge themselves to improve their reaction reflex at the expense of lowering accuracy. That behavior is validated in this study, as showed in Figure 6. Users that are more conservative obtain better accuracy, while those that score longer anticipation time tend to miss-answer in some of the sequences.

4) Misses: Errors on detection are mainly caused by thinking too far ahead in NLC sequences. When the target vehicle moves a little within the lane, the user tends to hit the button, labeling it as a LLC/RLC, leading up to an average of 1.85 errors per user (SD = 1.59). In LLC sequences, the error average is 0.60 (SD = 0.78), the same as on RLC sequences is 0.60 (SD = 1.02).

B. Participants

A total of 72 participants completed the survey, watching every selected sequence (30 out of 794). The average age of the participants was 24.22 years (SD = 10.08), and 74% of them were male. Most of them had a driving license (79%), and a 49% were daily drivers in highway scenarios. In Figure 7 the stats about the participants are detailed.

C. Statistical analysis

Null hypothesis states that the mean of delay to lane change starting point is 0, that is, participants can detect maneuvers precisely without any lag.

Performing a Student's t-test with a significance value of 1%, the null hypothesis is rejected, with a p-value of $p = 5.1615 \times 10^{-70}$ when testing that mean is not 0 and $p = 1.4184 \times 10^{-25}$ when testing as the alternative
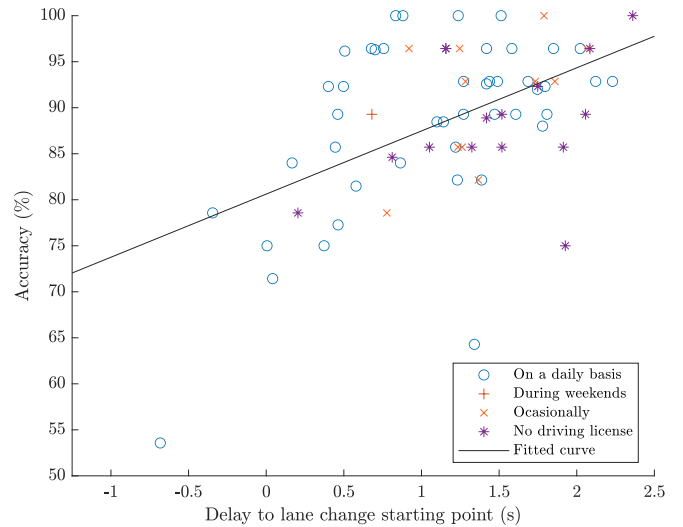

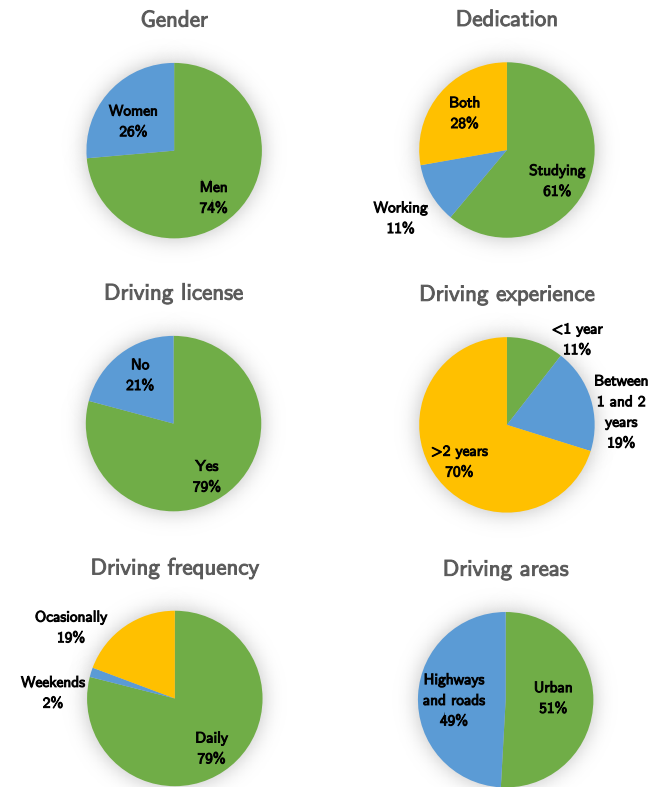Fig. 6. Delay-Accuracy ratio of participants ($r = 6.86x + 80.62$).

Fig. 7. Pie charts of users' data.

hypothesis that the mean is greater than 0. Therefore, the analysis revealed that **humans cannot anticipate a lane-change maneuver systematically**.

IV. DISCUSSION

A. Trial content and design

The survey was performed by researchers, looking for volunteers in the Polytechnic School where the INVETT Research Group is based. Moreover, professors and research colleagues were asked to take part in the Challenge, as well as students, and other assistant staff, along with individuals

related to the researchers (family, friends). The design of the Challenge has been limited to the resources available and keeping in mind the nature of personal supervision needed in every single trial. As said before, the maneuvers are slightly imbalanced (table I), being 47.61% of them RLC, 29.85% LLC, and 22.54% NLC. Nevertheless, it was determined to introduce the data in the same proportions for each test (33.33%) to make it impartial for the users. The dataset offers a reasonable amount of variability of all kinds of sequences, so considering the number of users and the random pick of them to create every single trial, this decision is adequate.

B. User groups

Risk-taking behavior has been analyzed within groups, being men the group that takes a more significant risk when making predictions of lane-change maneuvers. That is, they anticipate maneuvers better, but lowering the accuracy. As shown in Figure 6, participants that do not drive (i.e., no driving license) obtain the worst results in terms of delay and accuracy. Regarding driving experience, users whose experience is more than two years obtain the best results, especially in accuracy. Novice drivers score worse results than them. Considering driving areas, volunteers that are used to drive in highway scenarios outperformed the ones that are used to urban areas. The conclusions obtained in this analysis are consistent with the *a priori* judgment, especially in aspects such as driving experience and driving areas.

V. Conclusions and Future Work

72 participants took part in the study, trying to predict the direction of travel of a vehicle in 30 different sequences, selected from a set of up to 800. The vehicle selected in each sequence would perform a left lane change, a right lane change, or stay in the lane. On an overall level, participants were not able to predict the direction of the vehicle, compared with GT data. Certain factors, such as front wheels angling, merging maneuvers, or the use of turn signal increased participants' reliability, making them predict better the direction in those situations, as been qualitative observed.

As future work, the main task would be to improve the Challenge, adding cue labeling and extending the number of participants, even using an external survey participation service, such as Amazon Mechanical Turk. Adding more experienced drivers to the study would improve prediction correctness and certainty, which could lead to compare it to advanced prediction algorithms, such as [13], in order to develop better architectures and understand what is the best approach to predict lane changes successfully. Another exciting research work would be to monitor participants' eyes and faces with a frontal camera while completing the survey and analyze their reactions and cue extracting abilities. Finally, a future approach could be developed by using a simulator (e.g., CARLA [14]) along with an immersing cockpit that could assess users' capacity to avoid hazardous situations that they can spot while driving, being previously programmed in several trials regarding tailgating maneuvers, avoidance of rear-end collisions and evasive maneuvers when driving into an obstacle on the road.

References

[1] European Comission, "Road Safety In The European Union," vol. 29, no. 3, pp. 359–367, 2018. [Online]. Available: https://ec.europa.eu/roadsafety

[2] European Commision, "EU Road Safety Policy Framework 2021-2030 - Next steps towards "Vision Zero"," no. SWD(2019) 283 final, p. 33, 2019. [Online]. Available: www.etsc.eu/PIN

[3] S. S. Lee, E. E. Olsen, and W. W. Wierwille, "A Comprehensive Examination of Naturalistic Lane-Changes, Tech. Rep. DOT HS 809 702, 2004.

[4] S. Hetrick, "Examination of Driver Lane Change Behavior and the Potential Effectiveness of Warning Onset Rules for Lane Change or "Side" Crash Avoidance Systems," Ph.D. dissertation, 1997. [Online]. Available: http://scholar.lib.vt.edu/theses/available/etd-382915749731261/unrestricted/etd.pdf

[5] R. Izquierdo, A. Quintanar, I. Parra, D. Fernández-Llorca, and M. A. Sotelo, "The PREVENTION dataset: a novel benchmark for PREdiction of VEhicles iNTentIONs," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, Oct 2019, pp. 3114–3121.

[6] J. Nilsson, J. Silvlin, M. Brannstrom, E. Coelingh, and J. Fredriksson, "If, when, and how to perform lane change maneuvers on highways," *IEEE Intelligent Transportation Systems Magazine*, vol. 8, no. 4, pp. 68–78, winter 2016.

[7] R. Sukthankar, "Situational Awareness for Driving in Traffic A Thesis Proposal," Tech. Rep., 1994.

[8] G. H. Walker, N. A. Stanton, T. A. Kazi, P. M. Salmon, and D. P. Jenkins, "Does advanced driver training improve situational awareness?" *Applied Ergonomics*, vol. 40, no. 4, pp. 678–687, 2009.

[9] H. Zhao, C. Wang, Y. Lin, F. Guillemard, S. Geronimi, and F. Aioun, "On-Road Vehicle Trajectory Collection and Scene-Based Lane Change Analysis: Part i," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 1, pp. 192–205, jan 2017.

[10] S. De Craen, D. A. Twisk, M. P. Hagenzieker, H. Elffers, and K. A. Brookhuis, "Do young novice drivers overestimate their driving skills more than experienced drivers? Different methods lead to different conclusions," *Accident Analysis and Prevention*, vol. 43, no. 5, pp. 1660–1665, sep 2011.

[11] M. R. E. Romoser, D. L. Fisher, R. Mourant, J. Wachtel, and K. Sizov, "The Use of a Driving Simulator to Assess Senior Driver Performance: Increasing Situational Awareness Through Post-Drive One-on-One Advisement."

[12] E. Ohn-Bar, A. Tawari, S. Martin, and M. M. Trivedi, "Predicting Driver Maneuvers by Learning Holistic Features," in *IEEE Intelligent Vehicles Symposium, Proceedings*. Institute of Electrical and Electronics Engineers Inc., 2014, pp. 719–724.

[13] R. Izquierdo, A. Quintanar, I. Parra, D. Fernández-Llorca, and M. A. Sotelo, "Experimental validation of lane-change intention prediction methodologies based on CNN and LSTM," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, Oct 2019, pp. 3657–3662.

[14] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, " CARLA: An Open Urban Driving Simulator," in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1–16.