Driver state prediction from vehicle signals: An evaluation of segmentation approaches

Martin Maritsch[1], Kevin Koch[2], Hauke Thomsen[3], Niklas Kühl[3], Matthias Pfäffli[4], Wolfgang Weinmann[5], Felix Wortmann[2]

Abstract— **Modern vehicles typically are equipped with assistance systems to support drivers in staying vigilant. To assess the driver state, such systems usually split characteristic vehicle signals into smaller segments which are subsequently fed into algorithms to identify irregularities in driver behavior. In this paper, we compare four different approaches for vehicle signal segmentation to predict driver impairment on a dataset from a drunk driving study (n=31). First, we evaluate two static approaches which segment vehicle signals based on fixed time and distance lengths. Intuitively, such approaches are straightforward to implement and provide segments with a specific frequency. Next, we analyze two dynamic approaches that segment vehicle signals based on pre-defined thresholds and well-defined maneuvers. Although more sophisticated to define, the more specific characteristics of driving situations can potentially improve a driver state prediction model. Finally, we train machine learning models for drunk driving detection on vehicle signals segmented by these four approaches. The maneuver-based approach detects impaired driving with a balanced accuracy of 68.73%, thereby outperforming time-based (67.20%), distance-based (65.66%), and threshold-based (61.53%) approaches in comparable settings. Therefore, our findings indicate that incorporating the driving context benefits the prediction of driver states.**

I. INTRODUCTION

In recent years, with the advent of driver assistance systems, research in reliably predicting driver states has increased significantly. The aim is not only to enhance the comfort of drivers [1], but especially to improve road safety [2] by preventing accidents related to driver impairment. Advances in in-vehicle computing capacities and sensor technologies further support these endeavors. Besides camera-based technologies [3], the focus is on privacy-preserving and non-intrusive approaches [4], [5]. As outlined in recent work, a large number of in-vehicle signals can be accessed via the controller area network (CAN-bus), which allows insights into the driving behavior [6], [7].

The research field around gathering driver insights from in-vehicle signals is extensive. It ranges from driver identification [4], [5], [8], over driving style recognition [9]–[11] to driver state prediction for safety-critical conditions, such as drowsiness [12]–[14], drunkenness [2], [15]–[17], and more recently also hypoglycemia warning systems for drivers with diabetes [18]. All these applications use a common approach, they segment the high-frequency—and commonly also high-dimensional—vehicle signal data and subsequently apply descriptive statistics to represent these segments in feature vectors. Machine learning models can then leverage these feature vectors as input to reliably predict the driver state. The advantage of this approach is that it offers intelligible insights due to the interpretability of statistical features and, in addition, only requires slight amounts of computational load [6], [19].

The most common segmentation approach is to segment data based on a *fixed time sliding window (FTW)*. This approach is simple to implement but nonetheless achieves comparably good classification performance, for example, in identifying a driver [8], [20], detecting drowsiness [14], or estimating drivers' emotions [1].

Sathyanarayana et al. [21] propose a slightly adapted version of the FTW approach: instead of segmenting signals based on the time passed, they segment based on the distance traveled by a vehicle, therefore referring to it as *fixed distance sliding window (FDW)* approach. Compared to the FTW approach, the authors argue, that even at slow vehicle velocities the FDW approach yields a high information value of a segment because of the constant window distance.

Another widely used approach is to generate segments based on vehicle signals exceeding pre-defined *thresholds* [9], [10]. The popularity of this approach stems from the fact that driving maneuvers such as accelerating or braking change significantly, e.g., under the influence of alcohol [22], and, therefore, offer valuable insights into the driver state.

Similarly, *maneuver-based* approaches segment vehicle signals based on specific driving maneuvers inferred from vehicle location [2], [6]. This approach allows for an evaluation of driver behavior in similar situations, regardless of whether a particular signal threshold is exceeded or not. Hallac et al. [6] also point out the need to identify relevant situations to make prediction models more robust to signal noise. Further, findings of Lee et al. [15] indicate that the performance of driver state prediction depends on the driving maneuver, thus, arguing that the driving context plays a key role in algorithm development. Likewise, Li et al. [17] conclude that encoding information about road characteristics (i.e., radius and direction of a curve) is beneficial and training machine learning models for each curve type even increases

[1]Department of Management, Technology and Economics, ETH Zurich, Switzerland mmaritsch@ethz.ch

[2]Institute of Technology Management, University of St. Gallen, Switzerland {kevin.koch, felix.wortmann}@unisg.ch

[3]Institute of Information Systems and Marketing, Karlsruhe Institute of Technology, Germany hauke.thomsen@student.kit.edu, niklas.kuehl@kit.edu

[4]Department of Traffic Sciences, Institute of Forensic Medicine, University of Bern, Switzerland matthias.pfaeffli@irm.unibe.ch

[5]Forensic Toxicology and Chemistry, Institute of Forensic Medicine, University of Bern, Switzerland wolfgang.weinmann@irm.unibe.ch

the performance.

While these different segmentation approaches are well studied on their own, a comprehensive comparison with a standardized driving dataset is lacking. Furthermore, the analyzed datasets are frequently limited in the duration of driving and only encompass a small number of different driving situations, for example, by only driving in one environment (see [2], [14]) or having no interactions with other road users [23], making it difficult to compare the reported results.

To close this research gap, we apply a consistent setting to evaluate four commonly used segmentation approaches, including two sliding window approaches for fixed time and distance windows, as well as a threshold-based and a maneuver-based approach. As an illustrative dataset, we use vehicle data from a simulator study involving sober vs. drunk driving. Although the negative consequences of drunk driving are known, related accident rates are still high. In the US, 28% of fatal accidents involved drunk driving, resulting in 30 deaths per day on average [24]. The negative effects caused by alcohol that lead to these accidents are confirmed by a large body of literature [25]. Therefore, we consider the impairment of drivers due to alcohol an important and opportune use case for driver state prediction from vehicle signals and as a reasonable step to validate these segmentation approaches.

II. DATASET DESCRIPTION & METHODOLOGY

To test and compare the different segmentation approaches, we used a dataset that was collected as part of a study involving drunk driving. Where applicable, data is reported as mean $\pm$ standard deviation.

A. *Data acquisition*

The non-randomized, controlled, interventional single-center study included 31 participants (16 male, 15 female, age 37.65 ± 9.69 years) and was registered on ClinicalTrials.gov (NCT04980846). Study inclusion criteria were a valid driving license and an active driving record for the past two years, as well as a moderate level of alcohol consumption. Potential participants were excluded in case of health concerns or when taking illegal drugs or medications with known adverse effects in combination with alcohol.

Study participants were driving a passenger vehicle in a research-grade driving simulator (Carnetsoft BV, Groningen, The Netherlands). The driving routes featured a diverse set of intersection types (e.g., stop signs, traffic lights, and crosswalks) and included other road users and pedestrians to simulate naturalistic driving conditions. To cover a variety of driving situations, we implemented three different driving scenarios with distinct characteristics: highway, rural, and urban. These are as follows: (1) The highway scenario comprised of a two-lane highway with one-way traffic. Here, the route was mostly straight with a few wide curves. The speed limit was 80–120 km/h. Drivers experience varying traffic density, ranging from free flow to slow-moving traffic. (2) The rural scenario consisted of two-lane rural roads with traffic in both directions and several intersections with and without yield signs. The speed limit was 60–100 km/h. Drivers experienced other traffic participants and had to react to occasional events, such as a stopping bus or slower speeds in front of a school. (3) The urban scenario was used to reflect driving in a city. The route consisted of shorter and narrower roads compared to the two other scenarios. In addition, there were a large number of warning signs, intersections with and without yield or stop signs, and special events, such as pedestrians crossing streets. The speed limit was 30–50 km/h. Participants were instructed to adhere to local traffic laws (in Switzerland), act as they would in normal road traffic, and make use of all provided vehicle facilities, e.g., turn signal lights. While driving, participants had to follow the guidance of a navigation system. The study protocol included a pre-visit for individuals to familiarize themselves with the procedures, specifically with the simulator, and to rule out motion sickness.

The main study visit consisted of two driving sessions with different levels of controlled blood alcohol concentration (BAC): *sober* (0.00 g/dL BAC) and *drunk* (0.07 g/dL BAC), which is slightly above the legal limit of 0.05 g/dL BAC in most European countries. Each driving session for the two states consisted of driving for a total of 30 minutes ($\sim$29.7 km), with 10 minutes in each scenario (urban $\sim$3.7 km, rural $\sim$9.9 km, highway $\sim$16.1 km). Driving routes were identical across driver states and participants.

While driving, a multitude of signals capturing driver behavior and vehicle dynamics were sampled from an interface resembling a vehicle CAN-bus with a frequency of 30 Hz. In the present analysis, we use 7 vehicle signals as outlined in Table I. We additionally included any applicable speed limit from the driving context to assess vehicle velocity deviations from the allowed limits.

TABLE I: Extracted signals from CAN-bus

Driver behavior
Accelerator pedal position
Brake pedal position
Steering wheel angle
Vehicle behavior
Long. & lat. velocity
Long. & lat. acceleration
Speed limit deviation

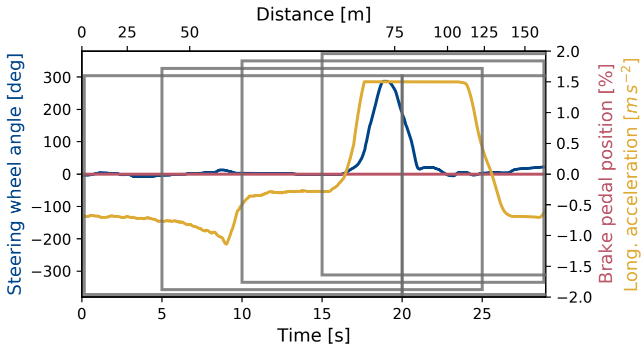
TABLE II: Descriptive statistics used to derive features

Derived features
min, max, sum, energy, mean, standard deviation, robust min (5th percentile), robust max (95th percentile), number of values above mean, number of values below mean, number of sign changes, square root of mean square

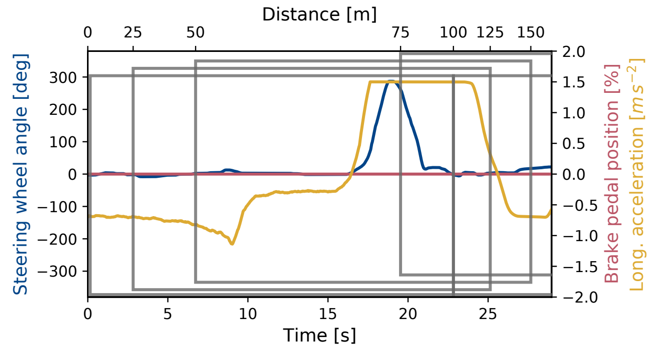
B. *Preprocessing*

The preprocessing of the extracted signals consisted of dropping the first 10 s and last 15 s of each trip to account for engine start and stop times. We also dropped sequences where the vehicle was standing still. In addition, we excluded sequences where drivers crashed or made wrong turns. In such cases, the drivers were reset to a prior position and had to start from the roadside, causing irregularities in the data.

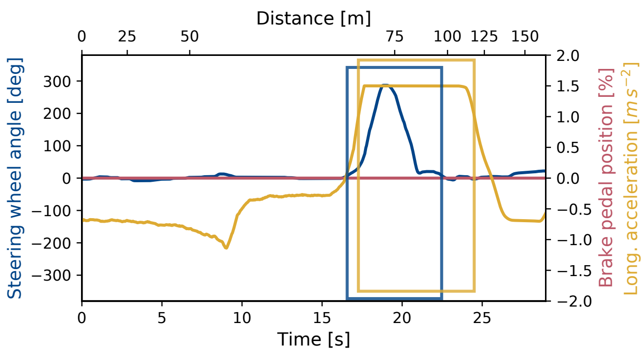
To enlarge the information captured from vehicle signal features, we included first (velocity) and second (accelera-



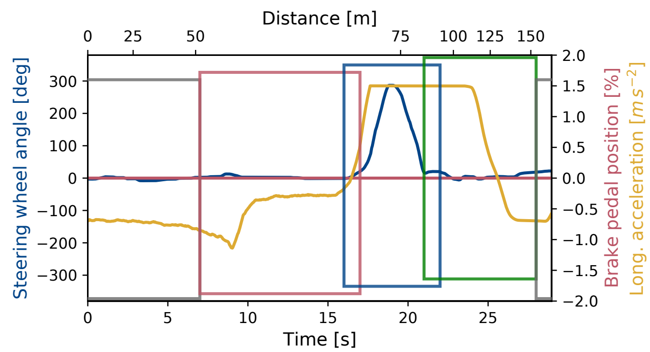
(a) Fixed time sliding window approach (20 s duration, 5 s step size) resulting in a new segment (gray) every 5 s.



(b) Fixed distance sliding window approach (100 m distance, 25 m step size) resulting in a new segment (gray) every 25 m.



(c) Threshold-based approach resulting in one steer (blue) and one accelerate (yellow) segment.



(d) Maneuver-based approach resulting in two straight (gray), one brake (red), one turn (blue), and one post turn (green) segment.

Fig. 1: Comparison of the resulting segments for the four segmentation approaches for an exemplary left turn at a stop sign.

tion) derivatives of steering wheel angle as well as accelerator and brake pedal positions.

C. Segmentation approaches

As outlined in Section I, high-dimensional signal data is commonly divided into segments in order to derive meaningful features that can be used as input to a machine learning model. In this paper, we analyze four different segmentation approaches and investigate their impact on the prediction performance for detecting drunk driving.

Fixed time sliding window (FTW). FTWs split the signal data into segments based on a fixed time duration (Fig. 1a). To allow for an extensive comparison, we analyze different window lengths, i.e., various time durations in the range of 5–60 s, similar to [14]. We applied a step size of 25% of the window duration so that two succeeding windows overlap by 75%. Since in our case vehicle signals are sampled at a constant frequency of 30 Hz, each window contains the same number of input data points.

Fixed distance sliding window (FDW). While being very similar to the FTW approach, the window length of the FDWs are not determined by the time duration but instead by the traveled distance (Fig. 1b). Again, we investigate a set of window lengths, i.e., distances, ranging from 100–1000 m. Fig. 2 shows that our chosen distances of FDWs result on average to window durations comparable to the FTW approach. Analogously, we applied a step size of 25% of the window distance. In comparison to the FTW approach, the number of data points in the FDW approach per window is variable since the window duration of the FDWs is inversely proportional to the current velocity.

Threshold-based segmentation. Compared to the previously discussed approaches, the threshold-based segmentation approach is not statically separating vehicle data into windows. Instead, segments are formed by inspecting the vehicle signals and isolating driving events based on predefined rules. Event beginnings, i.e., segment beginnings, are marked once the value of a signal exceeds a certain threshold. When the value drops below the threshold again, the endpoint of the corresponding segment is marked. Following previous work [9], [10], we focus on the three signals and associated thresholds that reflect the main vehicle controls used by a driver: *accelerating*, *braking*, and *steering*. We chose the following thresholds based on this previous work to identify driving events:

- Longitudinal acceleration $\geq 1.0\,ms^{-2}$
- Brake pedal position $\geq 1.0\%$
- Steering wheel angle $\geq 7.0\,\text{deg}$ or $\leq -7.0\,\text{deg}$

For an exemplary left turn at an intersection as depicted in Fig. 1c, the approach identifies two distinct events: an acceleration event and a steering event. Since the brake pedal is not used, no brake event is detected in this situation.

In a concluding step, we combined consecutive segments

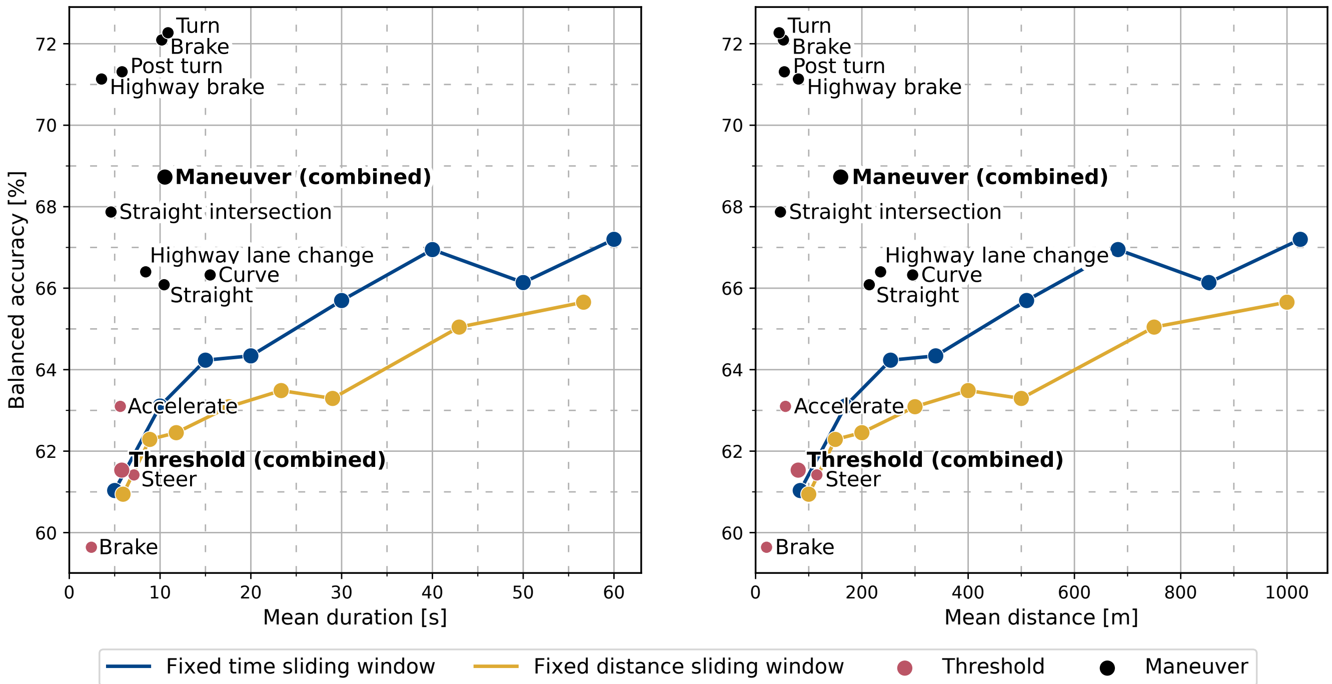


Fig. 2: Balanced accuracy of the different segmentation approaches over the mean segment duration in seconds (left) and mean segment distance in meters (right).

of the same event type if they are within 1 s of each other, assuming that they belong to the same driving intention of the driver. Further, we dropped all segments shorter than 1 s as we presume they contain too little information for the model to make a reliable prediction.

Maneuver-based segmentation. For the maneuver-based segmentation approach, we postulate that a modern vehicle is aware of its driving context. On real roads, GPS location and/or camera/radar systems can infer these information. Based on the driving routes in our dataset, we defined areas in which specific maneuvers occur (e.g., turning left at an intersection or braking before a stop sign). Similar to the threshold-based approach, we then determined the start and endpoints for each maneuver and driver, i.e., the locations where a driver started to brake or began turning the steering wheel. We calculated maneuver locations for each participant and subsequently obtained a common area over all participants, so that the trips of all drivers are covered within that specific maneuver. In comparison to the threshold-based approach, these maneuvers are independent of specific driver behavior and only rely on the layout of the road. For example, the maneuver-based approach captures the driver behavior in front of a stop sign and segments a braking maneuver as shown in Fig. 1d, even if the driver fails to brake at all.

We define and apply the following main maneuver groups based on the location, with their respective subgroups in parentheses:

- **Straight** ahead (urban, rural, highway)
- **Curve** (left, right)
- **Straight** at **intersections**
- **Turn** at intersections (left, right)
- **Post turn** after intersections (left, right)
- **Brake** (straight, left, right, speed bump)

In addition, we included the following maneuvers, which occur independently from absolute vehicle position, based on vehicle signal thresholds:

- **Highway braking**
- **Highway lane change** (left, right)

D. Feature generation

We aggregate the vehicle signals for each segment using descriptive statistics, resulting in statistical vectors which reflect the driver behavior for this segment. This approach is commonly used in related work [6], [19]. For each of the vehicle signals in Table I, including the derivatives of the driver behavior signals, we computed the 12 statistical features outlined in Table II [1], [2], [9]. Further, we added the time duration for each segment to the feature vector. In total, this process results in a 169-dimensional feature vector.

For the maneuver-based approach, we also incorporated information about the road context as one-hot encoded vectors. More specifically, we include information about the scenario (urban, rural, and highway) and, if applicable, the type of intersection (traffic light, stop sign, main road, etc.).

E. Predictive modeling

We used a decision tree-based gradient boosting machine learning model (LightGBM [26]) in its default configuration for the binary classification of the driver state (sober and drunk). We trained separate models for each window length

TABLE III: Results for the four segmentation approaches in detecting drunk driving. Fixed properties of sliding window approaches are underlined.

	Duration [s]	Distance [m]	Trip coverage [%]	Balanced accuracy [%]
Fixed time sliding window (FTW)	<u>5</u>	84.37 ± 46.49	100	61.03
	<u>10</u>	169.15 ± 91.24	100	63.11
	<u>15</u>	254.14 ± 134.62	100	64.23
	<u>20</u>	339.45 ± 177.10	100	64.34
	<u>30</u>	510.90 ± 261.25	100	65.70
	<u>40</u>	682.46 ± 344.45	100	66.95
	<u>50</u>	853.69 ± 425.97	100	66.14
	<u>60</u>	1025.15 ± 506.42	100	67.20
Fixed distance sliding window (FDW)	5.92 ± 4.20	<u>100</u>	100	60.94
	8.86 ± 5.96	<u>150</u>	100	62.29
	11.78 ± 7.70	<u>200</u>	100	62.45
	17.60 ± 11.08	<u>300</u>	100	63.09
	23.34 ± 14.45	<u>400</u>	100	63.49
	29.01 ± 17.66	<u>500</u>	100	63.29
	42.94 ± 25.26	<u>750</u>	100	65.04
	56.65 ± 32.87	<u>1000</u>	100	65.66
Threshold	$\mathbf{5.78 \pm 5.36}$	$\mathbf{80.17 \pm 111.79}$	$\mathbf{56.70 \pm 3.42}$	**61.53**
Accelerate	5.62 ± 3.54	56.24 ± 51.34	18.03 ± 3.44	63.10
Brake	2.43 ± 1.18	20.64 ± 16.95	5.51 ± 1.18	59.64
Steer	7.12 ± 6.46	115.30 ± 139.54	42.79 ± 2.11	61.41
Maneuver	$\mathbf{10.53 \pm 6.15}$	$\mathbf{160.16 \pm 165.25}$	$\mathbf{86.37 \pm 2.38}$	**68.73**
Straight	10.45 ± 7.18	213.91 ± 194.72	17.62 ± 3.75	66.08
Curve	15.52 ± 6.45	295.56 ± 186.69	35.45 ± 1.35	66.32
Straight intersection	4.60 ± 2.90	46.76 ± 11.48	2.47 ± 0.54	67.87
Turn	10.88 ± 3.22	44.29 ± 7.37	13.99 ± 1.80	72.27
Post turn	5.82 ± 2.05	54.12 ± 15.76	5.77 ± 0.39	71.31
Brake	10.19 ± 4.50	52.40 ± 25.49	21.55 ± 1.22	72.09
Highway brake	3.55 ± 0.70	80.59 ± 23.58	0.85 ± 0.61	71.13
Highway lane change	8.42 ± 1.86	235.27 ± 61.81	7.43 ± 2.40	66.40

of the FDW and FTW approaches. In the case of the threshold-based approach, we trained a separate model for each of the three driving events to allow the model to pick up peculiarities of that specific driving event [17]. For the maneuver-based approach, we proceeded analogously and trained a separate model for each subgroup of maneuvers.

We ran an iterative feature selection for each model to decrease the dimension of the input feature vector and, thus, also reduce overfitting of the model on the train data. The feature selection chose up to 50 features with the maximum positive impact on the overall prediction performance.

For model evaluation, we used leave-one-subject-out (LOSO) cross-validation, i.e., we trained the model on all segments of n–1 drivers and used the segments of the remaining driver as test data. This procedure was repeated for each participant, and the final results were averaged over all participants. The driving data used in this paper were obtained from a controlled setting in which the same routes are covered in each driver state. We evaluate the machine learning performance with balanced accuracy, therewith accounting for potential imbalances in single maneuvers and equally penalizing false positives and false negatives.

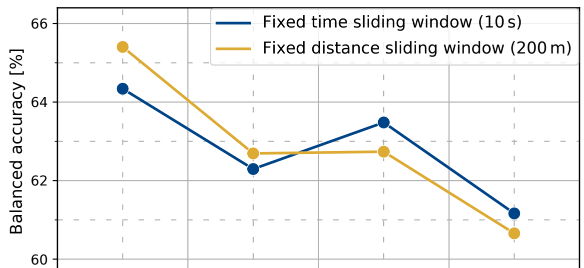
III. EVALUATION & RESULTS

Table III shows the driver state prediction results based on the four different segmentation approaches. In the following, we outline the results of the individual approaches in more detail and compare them. Since FTWs are the most commonly used segmentation approach, we use them as the baseline performance.

Fixed time sliding window. As previously described, the number of aggregated data points for calculating the feature vector is constant with the window length. Therefore, with increasing the window length, more information of the driving behavior and, thus, the driver state is captured. Our results show that the balanced accuracy improves with a longer time period covered, as depicted in the left panel in Fig. 2 and Table III. The balanced accuracy ranges from 61.03% (5 s) to 67.20% (60 s) with increasing performance for longer window durations. Since the time duration of the windows is fixed, the distance traveled by the vehicle within the time window depends on its velocity. For example, while 60 s of urban driving might only cover a turn at a busy intersection, 60 s of highway driving might cover a couple of hundred meters. With the sliding window approach, the entire driving trip is used for window segmentation, i.e., a trip coverage of 100% when taking all windows together, since the segmentation does not dynamically take the driving situation into account.

However, with increasing window length, data needs to be sampled for longer time periods before a prediction can be made. Hence, choosing the window length is a trade-off between the rapidity of the segmentation and the performance.

Further analysis of the results revealed that the balanced accuracy for segments with a lower mean velocity is considerably higher than for faster segments when keeping the window length constant (see blue line in Fig. 3, exemplary for FTW with 10 s). We attribute this difference to situations where the driver had to slow down in the urban scenario since such situations are likely more complex to handle, and driving deficits become more evident [27].



Velocity group [km/h]	< 30	30-60	60-90	> 90
FTW mean distance [m]	51.06 ± 19.42	123.32 ± 25.88	205.46 ± 21.54	308.58 ± 30.79
FDW mean duration [s]	32.07 ± 6.19	16.51 ± 3.66	9.72 ± 1.03	6.48 ± 0.67

Fig. 3: Balanced accuracy for the time-based (FTW 10 s) and distance-based (FDW 200 m) approach grouped by the mean velocity of the segments. Below, the mean distances (for FTW) and mean durations (for FDW) for each velocity group are listed.

Fixed distance sliding window. For the FDW approach, we observe similar results as for FTW. With increasing distance, the window becomes larger, and the balanced accuracy improves, ranging from 60.94% (100 m) to 65.66% (1000 m). Since FDW also statically segments vehicle signals into windows, it also achieves a trip coverage of 100%. However, again, increasing the window length is not always desirable, as an application might require the driver state to be predicted from a short segment [6].

Calculating the mean time duration of each window for the different distances allows a comparison with FTW. We can see that our defined window distances lead on average to similar long windows as defined by the FTW duration. Vice versa, the FDW distances are comparable to the mean distance of the FTW approach. Fig. 2 reveals that the balanced accuracy of the distance-based approach is always around 2% lower than that of the comparable FTWs.

When comparing both sliding window approaches for different velocities, then the distance-based approach outperforms the time-based one at lower velocities but loses its advantage at higher velocities. We attribute this to the fixed distance of the driving segment, for which the information value remains high even at low velocities where segments become longer in duration (see also [21]). In contrast, for FTW the segment duration remains constant, but the covered distance of a segment decreases with lower velocities, resulting in a lower informative value, which in turn has a negative impact on the performance. To illustrate this behavior, the mean window distances for FTW with 10 s and the mean durations for FDW with 200 m are shown for each velocity group in Fig. 3.

Threshold-based segmentation. Using thresholds to segment the signal data and training a model for each of the three different event types (i.e., accelerate, brake, and steer) results in balanced accuracies between 59.64% (brake) and 63.10% (accelerate) with a combined balanced accuracy of all event types of 61.53%. Acceleration events perform best, while brake events seem to distinguish only poorly between the two driver states. The lower balanced accuracy of brake segments can be explained at least in part by the comparatively short segment lengths of about 2.43 s, potentially resulting in insufficient informative value.

Noteworthy, the threshold-based approach only covers a somewhat limited portion of the driving trip. Merely in about 57% of the driving time one of the defined thresholds was exceeded, which means that the remaining 43% of the driving time was not taken into account for a driver state prediction. This comparatively small amount of segments shows the limitation of the threshold-based approach, more specifically, that a significant driver input/action is needed. On the other hand, with a mean duration of 5.78 s, the windows are relatively short and, therefore, allow for fast predictions in situations in which the driver needs to use the steering wheel or pedals frequently (e.g., within an urban scenario).

The combination of the three models for each driving event type is only slightly better than that of both sliding window approaches with comparable mean duration (i.e., FTW 5 s) and distance (i.e., FDW 100 m) (see Fig. 2). These findings are contradictory to related literature, as other work suggests that especially in these driving events, the differences between sober driving and drunk driving are significant [22], [28] and, thus, the performance should be better. We attribute the lower performance to our dataset, which includes a wide range of driving situations, making it difficult for machine learning models to compare such heterogeneous events, whereas related work primarily investigated specific driving events and situations.

Maneuver-based segmentation. The balanced accuracies for the different maneuvers range from 66.08% (straight) to 72.27% (turn), resulting in a combined balanced accuracy weighted by trip coverage of 68.73%. According to our analysis, turn maneuvers at intersections (72.27%), urban braking (72.09%), and post turn maneuvers (71.31%) reveal the driver state best. By separating the complex maneuvers based on the driving context and by training dedicated models, these models can better adapt to the behavioral changes in the alcohol-impaired state. These findings are in line with previous work [15], [17].

Although we segment the driving route into many smaller maneuvers, the proposed approach still covers 86.37% of the trip. Therefore, a prediction of the driver state can be performed in a fast and frequent manner with comparatively high balanced accuracy. In theory, implementing the maneuver-based approach would achieve 100% trip coverage, however, we had to discard some maneuvers during preprocessing in cases such as accidents.

Fig. 2 shows that the maneuver-based segmentation approach generates windows with similar short lengths as the threshold segments but by far outperforms its combined balanced accuracy of 61.53%. Moreover, our findings show that the maneuver-based approach outperforms both sliding window approaches with comparable window lengths by about 10%. The maneuver-based approach even outperforms FTW and FDW for considerably larger window lengths, highlighting the superiority of a segmentation approach respecting the driving context.

These promising results of the maneuver-based approach support our hypothesis that segmenting the driving route into distinct maneuvers allows for a more meaningful comparison of driver behavior in similar driving situations. By identifying high-level maneuvers not based on thresholds but rather from information about road features, the trained models can focus on the particular behavioral changes in specific driving situations. Therefore, machine learning models can accurately predict the driver state even for comparatively short road segments.

IV. Conclusion

This paper contributes to the ongoing research field of reliable driver state prediction from vehicle signals and highlights the impact of a chosen segmentation approach on the prediction performance. We investigated and compared four segmentation approaches (based on time, distance, threshold, and maneuver) on driving data of 31 participants from a study involving sober vs. drunk driving. The driving route consisted of three different scenarios (urban, rural, and highway) and, thus, covered a broad range of driving situations. Our results show that, while it is beneficial to have longer windows for the sliding window approaches (time and distance), the performance of driver state prediction improves when the driving route is segmented into distinctive driving maneuvers. Separating the specific driving situations allows machine learning models to better learn a representation of the behavioral changes in drivers under the influence of alcohol. In addition, the shorter maneuver-based segments are also better suited for providing timely predictions [6], compared to the significantly longer sliding windows with similar performance. Using shorter segments is particularly relevant considering that an intervening decision towards a specific driver state should be robust and, thus, such a system should consider several consecutive predictions, e.g., by majority vote [8], [20], before a final prediction is made.

Albeit the simulator setting allows for a controlled and balanced driving dataset, it also implies certain limitations. Although the study setting aimed to resemble driving scenarios as realistic as possible, the vehicle handling and driving conditions are still different from driving on real roads. However, related work indicates that alcohol-induced changes in driving behavior are comparable between simulator settings and real vehicles [29]. Another limitation is that automotive technology is rapidly evolving, and fully autonomous driving might resolve the issue of traffic accidents caused by driver impairment. Nevertheless, as of now, autonomous driving without the need for a vigilant driver (level 4 or 5) is deemed to still be several decades away due to increasing safety concerns associated with the underlying technology [30]. Therefore, solutions that bridge the intervening time by more rapidly and directly addressing the traffic incidents related to driver impairment are an important step towards safer roads.

While we report promising results for a more context-specific segmentation approach to detect impairment caused by alcohol, its transferability to the detection of other kinds of impairment remains to be verified by future work. For example, we anticipate a high potential of context-specific segmentation approaches for drowsiness detection since this condition causes similar impairment characteristics as alcohol, and related work already reported encouraging results for CAN data-based drowsiness detection [14]. A promising field of research lies ahead.

Acknowledgment

This work was supported by the Bosch IoT Lab, a collaboration of Bosch, ETH Zurich and the University of St. Gallen, Switzerland, by a fellowship within the IFI programme of the German Academic Exchange Service (DAAD), and by the Swiss National Science Foundation (Project CRSII5_183569).

References

[1] S. Liu, K. Koch, Z. Zhou, S. Föll, X. He, T. Menke, E. Fleisch, and F. Wortmann, "The Empathetic Car," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 3, pp. 1–34, 9 2021.

[2] Y. Yao, X. Zhao, H. Du, Y. Zhang, G. Zhang, and J. Rong, "Classification of fatigued and drunk driving based on decision tree methods: A simulator study," *International Journal of Environmental Research and Public Health*, vol. 16, no. 11, 2019.

[3] S. Liu, K. Koch, Z. Zhou, M. Maritsch, X. He, E. Fleisch, and F. Wortmann, "Towards Non-Intrusive Camera-Based Heart Rate Variability Estimation in the Car under Naturalistic Condition," *IEEE Internet of Things Journal*, 2021.

[4] J. Yang, R. Zhao, M. Zhu, D. Hallac, J. Sodnik, and J. Leskovec, "Driver2vec: Driver Identification from Automotive Data," in *MileTS '20: 6th KDDWorkshop on Mining and Learning from Time Series*, 2021.

[5] B. Gahr, S. Liu, K. Koch, F. Barata, A. Dahlinger, B. Ryder, E. Fleisch, and F. Wortmann, "Driver Identification via the Steering Wheel," *arXiv*, 2019.

[6] D. Hallac, A. Sharang, R. Stahlmann, A. Lamprecht, M. Huber, M. Roehder, R. Sosič, and J. Leskovec, "Driver identification using automobile sensor data from a single turn," in *IEEE Intelligent Transportation Systems*, 2016, pp. 953–958.

[7] U. Fugiglando, E. Massaro, P. Santi, S. Milardo, K. Abida, R. Stahlmann, F. Netter, and C. Ratti, "Driving Behavior Analysis through CAN Bus Data in an Uncontrolled Environment," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, 2019, pp. 737–748.

[8] B. Gahr, B. Ryder, A. Dahlinger, and F. Wortmann, "Driver Identification via Brake Pedal Signals - A Replication and Advancement of Existing Techniques," in *IEEE International Conference on Intelligent Transportation Systems*, vol. 21, 2018, pp. 1415–1420.

[9] P. Brombacher, J. Masino, M. Frey, and F. Gauterin, "Driving event detection and driving style classification using artificial neural networks," in *IEEE International Conference on Industrial Technology*, 2017, pp. 997–1002.

[10] M. Van Ly, S. Martin, and M. M. Trivedi, "Driver classification and driving style recognition using inertial sensors," in *IEEE Intelligent Vehicles Symposium*, 2013, pp. 1040–1045.

[11] Z. Kenkar and S. AlHalawani, "Event-Based Driving Style Analysis," in *Advances in Data Science, Cyber Security and IT Applications*, A. Alfaries, H. Mengash, A. Yasar, and E. Shakshuki, Eds. Communications in Computer and Information Science, Springer, 2019, vol. 1097, pp. 170–182.

[12] A. D. McDonald, J. D. Lee, C. Schwarz, and T. L. Brown, "A contextual and temporal algorithm for driver drowsiness detection," *Accident Analysis and Prevention*, vol. 113, pp. 25–37, 2018.

[13] A. Eskandarian and A. Mortazavi, "Evaluation of a smart algorithm for commercial vehicle driver drowsiness detection," *IEEE Intelligent Vehicles Symposium*, pp. 553–559, 2007.

[14] M. S. Wang, N. T. Jeong, K. S. Kim, S. B. Choi, S. M. Yang, S. H. You, J. H. Lee, and M. W. Suh, "Drowsy behavior detection based on driving information," *International Journal of Automotive Technology*, vol. 17, no. 1, pp. 165–173, 2016.

[15] J. D. Lee, T. Brown, D. Fiorentino, J. Fell, E. Traube, and E. Nadöer, "Using vehicle-based sensors of driver behavior to detect alcohol impairment," in *22nd International Technical Conference on the Enhanced Safety of Vehicles (ESV)*, Washington, DC, 2011.

[16] Z. Li, X. Jin, and X. Zhao, "Drunk driving detection based on classification of multivariate time series," *Journal of Safety Research*, vol. 54, no. June, pp. 29–64, 2015.

[17] Z. Li, H. Wang, Y. Zhang, and X. Zhao, "Random forest–based feature selection and detection method for drunk driving recognition," *International Journal of Distributed Sensor Networks*, vol. 16, no. 2, 2020.

[18] V. Lehmann, M. Maritsch, T. Zueger, A. Marxer, C. Bérubé, M. Kraus, C. Albrecht, S. Feuerriegel, T. Kowatsch, E. Fleisch, F. Wortmann, and C. Stettler, "5-LB: A Machine Learning–Based Approach to Noninvasively Detect Hypoglycemia from Gaze Behavior While Driving," *Diabetes*, vol. 70, no. Supplement_1, 2021.

[19] J. Xie, A. R. Hilal, and D. Kulic, "Driving Maneuver Classification: A Comparison of Feature Extraction Methods," *IEEE Sensors Journal*, vol. 18, no. 12, pp. 4777–4784, 2018.

[30] P. Bansal and K. M. Kockelman, "Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies,"

[20] M. Enev, A. Takakuwa, K. Koscher, and T. Kohno, "Automobile Driver Fingerprinting," *Proceedings on Privacy Enhancing Technologies*, vol. 1, pp. 34–51, 2016.

[21] A. Sathyanarayana, N. Krishnamurthy, and J. H. Hansen, "Analysis of driving maneuvers: Is the secret in the distance or time?" *SAE Technical Papers 2014-01-0299*, 2014.

[22] A. K. Yadav and N. R. Velaga, "Effect of alcohol use on accelerating and braking behaviors of drivers," *Traffic Injury Prevention*, vol. 20, no. 4, pp. 353–358, 2019.

[23] C. Irwin, E. Iudakhina, B. Desbrow, and D. McCartney, "Effects of acute alcohol consumption on measures of simulated driving: A systematic review and meta-analysis," *Accident Analysis and Prevention*, vol. 102, pp. 248–266, 2017.

[24] "TRAFFIC SAFETY FACTS - Overview of Motor Vehicle Crashes in 2019," *US Department of Transportation, NHSTA*, 2019.

[25] S. Jongen, E. F. Vuurman, J. G. Ramaekers, and A. Vermeeren, "The sensitivity of laboratory tests assessing driving related skills to dose-related impairment of alcohol: A literature review," *Accident Analysis and Prevention*, vol. 89, pp. 31–48, 2016.

[26] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems*, Long Beach, CA, USA, 2017, pp. 3147–3155.

[27] T. L. Martin, P. A. Solbeck, D. J. Mayers, R. M. Langille, Y. Buczek, and M. R. Pelletier, "A review of alcohol-impaired driving: The role of blood alcohol concentration and complexity of the driving task," *Journal of Forensic Sciences*, vol. 58, no. 5, pp. 1238–1250, 2013.

[28] X. Zhao, X. Zhang, J. Rong, and J. Ma, "Identifying method of drunk driving based on driving behavior," *International Journal of Computational Intelligence Systems*, vol. 4, no. 3, pp. 361–369, 2011.

[29] A. Helland, G. D. Jenssen, L. E. Lervåg, A. A. Westin, T. Moen, K. Sakshaug, S. Lydersen, J. Mørland, and L. Slørdal, "Comparison of driving simulator performance with real driving after alcohol intake: A randomised, single blind, placebo-controlled, cross-over trial," *Accident Analysis and Prevention*, vol. 53, pp. 9–16, 2013.
Transportation Research Part A: Policy and Practice, vol. 95, pp. 49–63, 2017.