

Driver Identification via the Steering Wheel

Bernhard Gahr*, Shu Liu⁺, Kevin Koch*, Filipe Barata⁺, André Dahlinger*, Benjamin Ryder⁺, Elgar Fleisch⁺,
and Felix Wortmann*

Abstract—Driver identification has emerged as a vital research field, where both practitioners and researchers investigate the potential of driver identification to enable a personalized driving experience. Within recent years, a selection of studies have reported that individuals could be perfectly identified based on their driving behavior under controlled conditions. However, research investigating the potential of driver identification under naturalistic conditions claim accuracies only marginally higher than random guess. The paper at hand provides a comprehensive summary of the recent work, highlighting the main discrepancies in the design of the machine learning approaches, primarily the window length parameter that was considered. Key findings further indicate that the longitudinal vehicle control information is particularly useful for driver identification, leaving the research gap on the extent to which the lateral vehicle control can be used for reliable identification. Building upon existing work, we provide a novel approach for the design of the window length parameter that provides evidence that reliable driver identification can be achieved with data limited to the steering wheel only. The results and insights in this paper are based on data collected from the largest naturalistic driving study conducted in this field. Overall, a neural network based on GRUs was found to provide better identification performance than traditional methods, increasing the prediction accuracy from under 15% to over 65% for 15 drivers. When leveraging the full field study dataset, comprising 72 drivers, the accuracy of identification prediction of the approach improved a random guess approach by a factor of 25.

Index Terms—Driver Identification, GRU, Machine Learning, Neural Networks, Time Series Signals.

I. INTRODUCTION

Driver identification is considered as the first step towards analyzing driver behavior, including the realm of complex classifications of vehicle operation behavior. For example, detecting driver distraction or fatigue for safety enhancement features [1], or enabling systems that identify critical health issues such as heart or severe hypoglycemia attack. The frictionless identification of individual drivers enables various benefits, including a streamlined driving experience without the need for cumbersome authentication, improved security through intrusion detection systems, and personalized in-car service models [2]. Naturally, automated identification raises questions regarding privacy — especially due the increasing interconnectivity of modern cars [3] — and the recent work of multiple research groups continues to show growing interest in the potential of driver identification.

The high potential of utilizing CAN-bus data collected from the vehicle has been the main focus of recent research on

driver identification [4], [5]. For example, specialized analysis methods for a single CAN-bus signal can significantly improve identification accuracy [6], [7], particularly when considering the brake pedal and throttle signal. Some extreme results have claimed the ability to perfectly identify a driver [8], [9] despite using just one CAN-bus signal such as the brake pedal [3]. These results indicate the power of driver identification via longitudinal vehicle control. However, all studies on this topic come with a variety of limitations. These include study design issues, typically from collecting data under a controlled driving setting. Research based on naturalistic driving data has not been able to recreate the extreme results. Further, current research based on naturalistic driving data and relying on a multitude of sensors suffers from a potential misfitting of the trained model onto the car and not the driver [9]. Finally, research limited to a single sensor suffers from practical issues, such as the necessity to observe the driver for an unsuitably long time until reliable identification can be made [6]. Additionally, even low- and mid-range cars, such as the Volkswagen Golf or the Ford Fusion, currently come equipped with adaptive cruise control that takes over the longitudinal vehicle control. Therefore, signals from the brake and gas pedals will no longer reveal any strong information about the driver, rendering existing identification models and research obsolete. In order to advance driver identification and remain relevant in the coming years, new classification methods and strategies must be considered. Specifically, those signals whose inputs are likely to remain in the hands of the driver, i.e. lateral control of the vehicle from the steering wheel, are set to become the only reliable information sources for driver identification.

In this paper we introduce an approach that leverages the steering wheel for driver identification. Our research builds upon on an extensive literature review, which, to our best knowledge, does not yet exist for the topic of driver identification. We show that the existing methods utilized in previous work, typically assessed in controlled driving settings, could not deliver their reported accuracies when applied to real-world naturalistic driving data. Against this background, we present and evaluate a machine learning approach that performs well in this challenging setting, and which is motivated and based on statistical analyses of the steering wheel signal.

This paper is structured as follows: we initially present research related to the topic of personal identification via vehicle data and highlight the existing research gap. We then describe the study design and field test setting, followed by the analytical methodologies we applied. Further, we present results validating the reviewed methods under a naturalistic setting, then show how identification accuracy improves when

*Institute of Technology Management, University St. Gallen, Switzerland {bernhard.gahr, kevin.koch, andre.dahlinger, felix.wortmann}@unisg.ch

⁺Department of Management, Technology and Economics, ETH Zurich, Switzerland {liush, fbarata, bryder, efleisch}@ethz.ch

the presented novel feature collection method is applied. Finally, we conclude with a discussion of the results and the implications of this research.

II. RELATED WORK

This section provides an extensive literature review. We extended this section compared to other publications, because the topic of driver identification using CAN-bus data has become almost a trending topic in recent years. While the first research was done in 2005 [10], sixteen out of the twenty publications we found appeared in the last four years. However — to the best of our knowledge — there is no comprehensive literature research, leading to partially overlapping and partially contradictory published research. We first present research based on simulator and controlled field study data. In Section II-B we present research based on naturalistic driving data. A compact overview of the research is presented in Table I in the Appendix. For the sake of compactness we neglect a closer description of works that used sitting posture [11], finger vein [12] or voice [13] signals to identify a driver and only describe research that relied on data available on the CAN-bus.

A. Simulator- and Controlled Driving Studies

The first research on driver identification based on CAN-bus signals collected through a driving simulator led to the identification of drivers with an accuracy of up to 73% for 30 drivers [10]. To achieve this result, the authors consider driving behavior signals including the accelerator pedal, brake pedal, and vehicle velocity. Additionally they took following distance into account. Later, the same group was able to optimize re-identification to 89.6% for the simulator, which dropped to 71% for real car data, by applying spectral analysis methods [14]. In a study of 276 drivers measuring simulator- and real-car signals as well as the following distance, the group was able to re-identify a driver with 76.8% accuracy [15].

Almost ten years later, the next study focused on the implementation of neural nets of driver identifying methods [16]. The authors used data from a controlled field study [17], where drivers had to follow a predefined route of 25 km. Results were limited to detecting subsets of three to five of the available eleven drivers, reaching identification accuracies from 70% to 85% for different settings. In a second step [18] they used an extreme learning machine (ELM) network which was applied on audio-, video-, inertial measurement unit-, frontal laser scanner-, and CAN-bus signals. Identification accuracies increased for two to eleven drivers to a range of 85% to 97%. In the same year the first research group claimed perfect identification for a set of 15 drivers [3]. The claim was made even stronger by achieving perfect identification using only 15 minutes of driving data, or 90 minutes of driving data from the brake pedal only. Using only the steering wheel signal, the group achieved an accuracy of 83.33% for 90 minutes of training data. As in the previous research, the data was collected in a controlled field study. These results are often considered state-of-the-art, since many later publications based their work on them. In [19], a second group claimed perfect

identification not only after each session, but by continuously authenticating a driver throughout a driving session. Their results were based on data from a simulator study. Similar to [19], [8] focused on the immediate identification of a driver by using as few data as possible and claimed perfect identification of six to ten drivers using only five minutes of training data from controlled driving studies [20], [21], [22].

In [23] the researchers also focused on the immediate identification of a driver. However, similar to [3] they contributed with analysis on a sliding window design for a random forest (RF) algorithm. In contrast to the optimal 3-second window in [3], they found their optimal window length at around 15 seconds, with an accuracy of 89% for 15 and 82% for 35 drivers. In [24] the authors focused again on the real-time identification of a driver using convolutional neural networks (CNN). They argued that performances for larger sets drop using conventional RF, support vector machine (SVM) or ELM algorithms. With 75 seconds of training data and a 0.25 seconds sliding window design they achieve 88% accuracy for four drivers. In [7] the researchers focus on the design of SVMs and universal background models (UBM) to identify four drivers using only brake and throttle signals from a controlled driving study. They achieve an identification rate of 83% by taking an approximate window length of 5 seconds for the throttle and 15 seconds for the brake. In [25] the researchers aim to build an intrusion detection system based on CAN-bus values. For 10 drivers, recognition accuracies over 98% were achieved using extra trees (ET). In [26] the authors introduce a multi-layer perceptron network to identify up to ten drivers with an accuracy of 95% in a controlled field study. Additionally they investigate the accuracies of window sizes of 1, 10, 30, and 60 seconds. Contradicting previous publications, they conclude that a window size of one minute leads to the best result. In [27], based again on a controlled field study, the authors use a RF algorithm to identify a subset of 4 out of 15 drivers with an accuracy of 89%. The results are achieved by tuning the parameters for the number of trees and number of features.

B. Naturalistic Driving Studies

The first results using real naturalistic driving data were published in [28] and [4]. Using a 30-second window for feature extraction, the researchers were able to perfectly distinguish between two drivers [28]. Accuracies dropped when additional drivers were added to below 40% for nine and more. In [9], researchers approached the topic by applying findings and results from [3] onto the collected naturalistic driving data with an optimal window length of 5 seconds, which lead to perfect identification of 30 drivers. However, to achieve this, all available CAN-bus signals and a training set of over four hours and at least ten minutes of driving data for testing was required. Limitations, according to the authors, of leveraging such a data set were partially missing signals and potential fitting of the model to the cars' differences since each participant was using his/her own car. "However, the different mileage and ages of vehicles might cause differences in vehicle dynamics. That means the random forest model might potentially pick

up some differences between the vehicles.” [9]. Further results on naturalistic driving data is presented in [29]. The authors were able to identify 50 drivers with an accuracy of 40% by using 160 trips per driver for training and 40 trips for testing. For that, a special Autoencoder Regularized Network (ARNet) was developed consisting of stacked recurrent neural networks (RNN) using gated recurrent units (GRU) and fully connected (FC) layers with a window size of 4 seconds.

A fundamentally new approach for driver identification was taken in [4]. The authors limited themselves only to turning events and tried to identify the drivers leveraging all available CAN-bus signals. The approach showed promising accuracy improvements over a random guess, however with 50% accuracy for 5 drivers it did not reach the performance of previous results. In a later approach, the same research group increased identification rates to 51% for 56 drivers by applying GRUs with a 1-second window onto the available CAN-bus signals [5]. Similar to [9], an enormous amount of driving data (≈ 23 hours per driver) was required to train the model. Following the event-based approach of [4], the authors of [6] tried to identify up to 50 drivers by only using the brake pedal signal. They limited themselves to the brake pedal signal, collected under a naturalistic driving setting, based on the results in [3] and showed that the brake pedal indeed serves as a good identifier of the driver. Additionally they showed that identification performances can be significantly increased to almost perfect identification for 5 and 85% accuracy for 50 drivers by tailoring a model to one specific signal. However, they also showed that the brake pedal is only sparsely used over time while driving.

Concluding, perfect driver identification has only been achieved under controlled data collection settings. High accuracies were achieved using sensor data from the longitudinal vehicle control (i.e. brake pedal and throttle). The reported window sizes for the models are contradictory in the literature and range from under one second to over a minute. Using naturalistic driving data, two key findings were identified from related work. First, an enormous amount of data is required to achieve sufficiently good results. And second, if every study participant uses their own car, limiting to a single sensor also limits the chance of misfitting the model onto the car. Therefore we identified the research gap of driver identification using naturalistic driving data with the limitation on lateral vehicle control (i.e. steering wheel signal).

III. DATA & METHODS

In the following, we describe our signal processing methods and underlying theoretical approach. The analysis is based on data collected from a field study under naturalistic driving conditions of 72 patrollers from a road assistance service. The patrollers were distributed over 10 different base locations in the German-speaking part of Switzerland, covering mountainous as well as hilly to almost flat regions. They were not asked to perform any special maneuvers and followed their daily routine. All but one driver was male, ranging in age from 21 to 64 years old with a mean of 40 years. Further, every participant was driving a Chevrolet Captiva. The study

ran for approximately five months. In total, over 680'000 km of driving data was collected by reading out CAN-bus data via the ODB-II port with a dongle that transferred the signals via Bluetooth to a smartphone, where GPS and accelerometer signals were also added and then transmitted in real time via the cellular network to a server. The 72 drivers transmitted on average over 52 hours of driving data.

A. Preprocessing

Each trip is queried separately from the database containing the raw collected data. This means that the recorded data is prone to the inconsistent sampling, Bluetooth data transmission errors from the dongle to the smartphone, and transmission outages from the smartphone via the cellular network to the server. The trips were queried as a *Pandas DataFrame*¹. We first cleaned the data by removing all erroneous queried values, such as *NaNs*. Additionally, all data points with GPS outage were removed.

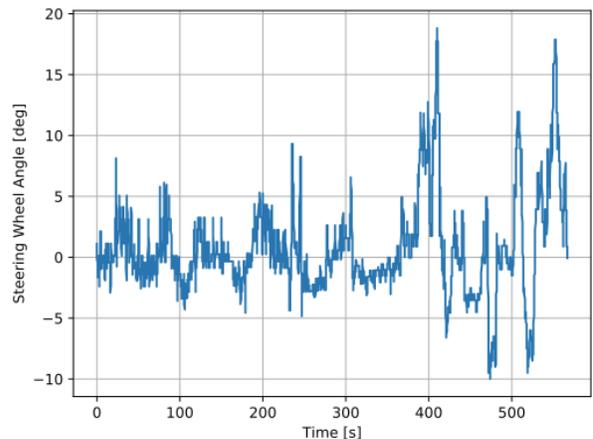


Fig. 1: Example of a recorded steering wheel signal.

In a second step, the data is resampled and quadratically interpolated at a frequency of 10 Hz, consistent with prior research [3]. Last, each of the trips is required to have a minimum length of 5 minutes. When a preprocessed trip is shorter than the required minimum length, the trip is dismissed. An example of a recorded and preprocessed signal can be seen in Figure 1. In the following, we describe the signal of each trip by $X_{i,t}$ where $t \in [0, T_i]$ and T_i denotes the total number of samples of the i -th recorded trip.

B. Stationarity & Correlated Lag

We motivate the design of our neural net by the analysis performed on the collected steering wheel signals in this section. In the following we will show that the steering behavior can be modeled as a stationary signal. This will allow us to draw concrete design decisions for an optimal window length for the models introduced in Section III-C2.

¹<https://pandas.pydata.org>

For simplicity, we neglect the trip notation and simply write X_t for an observed signal. Following the description of [30], the expected value of the observed signal is denoted by $E\{X\}$ and the covariance of X_t with the signal itself at a different time step X_{t+h} is denoted as $Cov\{X_t, X_{t+h}\} = E\{X_t X_{t+h}\} = \gamma(t, h)$. In the following we use the definition of [30] of a stationary time series:

- 1) the expected value is a constant for all t , i.e. $E\{X_t\} = c, \forall t$, and
- 2) the covariance is time-independent, hence only depends on the lag h , or time distance, between two values, i.e. $Cov\{X_t, X_{t+h}\} = E\{X_t X_{t+h}\} = \gamma(h), \forall t$.

Last, we introduce the autocorrelation function

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \quad (1)$$

which normalizes the autocovariance function, s.t. the following properties are valid: $\rho(0) = 1$, $|\rho(h)| \leq 1$, and $\rho(h) = \rho(-h)$.

Following [30], we test for a unit root at a 1% significance level of all trips². In case the H0 hypothesis of a unit root is rejected, i.e. a p-value below the 1% significance level is returned, we assume the steering wheel signal to be stationary. Our analysis showed that over 12'278 tested trips of all 13'883 were returned with a rejected H0 hypothesis. Using the stationary property of a process without a constant frequency, we can further assume a fading ACF, i.e. after some lag $t + h > t + h_{cor}$, the observed measurements are not significantly correlated with the the measured value at t . The property of a fading ACF does also hold for the 1'605 signals where the H0 hypothesis was not rejected, however finding the uncorrelated lag h_{cor} is extremely cumbersome since Equation (1) becomes time-dependent and hence the lag has to be found at each sampling step.

Due to the symmetry property, it is sufficient to only look at the right-sided ACF. An example of the right-sided ACF of a recorded steering wheel signal is shown in Figure 2. In the plot we can see that the correlation of a signal value at time lag h fades out and stays within a significantly uncorrelated region, as the lag h gets larger. The lag at which the ACF and the significance region intersect (h_{cor}) is displayed in Figure 2 by the vertical red line, while the significance region is shown in light blue.

For the design of our neural net, we want to select a window size that corresponds to a \hat{h}_{opt} which is large enough so that all significantly correlated values with lag $h_{cor,i}$ are still taken into account. At the same time, we want to avoid \hat{h}_{opt} being too large, so that uncorrelated (noise-contributing) values are taken into account. Following this approach, we analyzed all stationary trips. Doing that, we limited the maximum lag to 20 seconds. For the analysis we required again a significance level of 1%. Figure 3 shows the histogram of all found lags. The mode of the histogram lies at 3.6 seconds, while the median and mean values lie at 5.8 seconds and 6.6 seconds, respectively. This indicates that an optimal window length

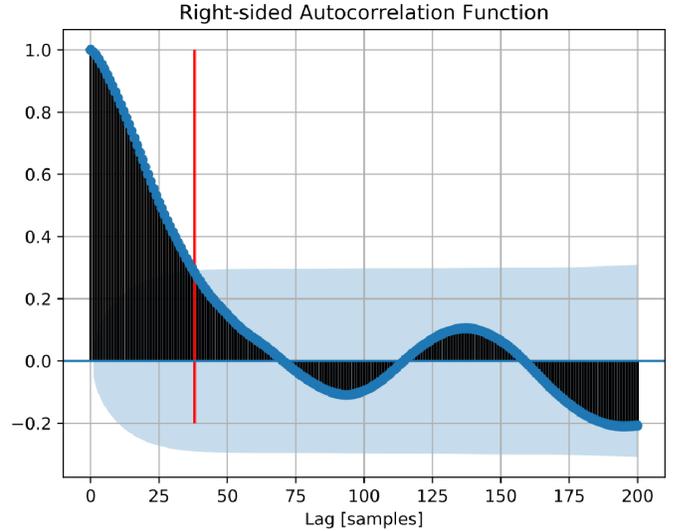


Fig. 2: Example of a right-sided autocorrelation function and the 1% confidence region.

should also lie in the range of 2 to 10 seconds. In Chapter IV-A we validate these findings and present an optimal window size \hat{h}_{opt} .

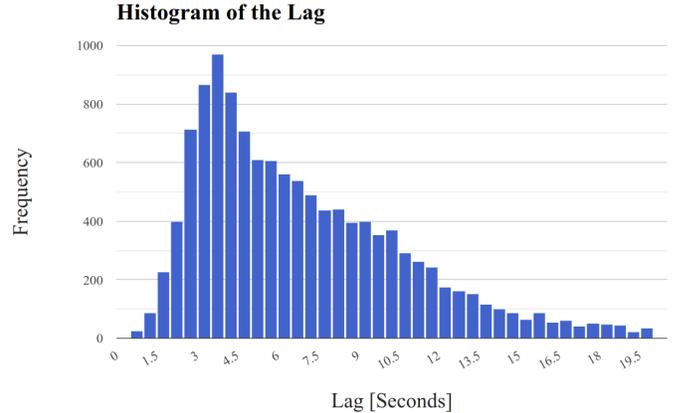


Fig. 3: Distributions of the maximum lags for all stationary trips.

C. Method Description

In the following we describe the signal's postprocessing and feeding into the neural network. In contrast to previous attempts using GRUs [5], we do not rely on the raw signals but compute the logarithmic spectrogram as additional features and feed them into the network to improve the accuracy [31].

1) *Feature Preparation:* We first extract segments of a predefined length from each trip. Since the trips might vary in length, this operation helps ensure a fixed length. In our case, we define the length of a segment to be $S = 15$ min. Multiple segments can be extracted from one trip, if $T_i > S$, and one segment can contain multiple trips, if $S > T_i$, by simply concatenating the trips. We make sure every driver has the same number of segments. This way, we ensure a balanced

²We used the library <http://www.statsmodels.org> in its version 0.9.0 for python3

dataset for training and testing. For each segment, we apply a sliding window \vec{W}_w without overlapping to extract features. The length of the sliding window relates the previously found \hat{h}_{opt} . We denote these windows as $\vec{W}_w[t]$ with $t \in [0, \hat{h}_{opt}]$. Each of the windows is then used to calculate one feature vector \vec{F}_w . We chose a rather simple feature representation of the signal and only took the logarithmic Fourier-Transformation (LFT) of a window:

$$\vec{F}_w = \log_2(|\text{FFT}(\vec{W}_w)| + 1), \quad (2)$$

where $\text{FFT}(\cdot)$ denotes the Fast-Fourier-Transformation and hence, \vec{F}_w is of length \hat{h}_{opt} .

Additionally to the LFT, the average velocity of the car in each window was appended to \vec{F}_w . The motivation for the velocity is twofold. First, the transmission from the steering wheel angle to the actual wheel angle — and hence the curve radius — is nonlinear in modern cars. To avoid dangerous strong steering at high speeds, which could lead to a potential rollover of the car, the steering wheel has a lower angle transmission towards the wheels at those speeds [32]. Second, the centripetal force is linearly dependent on the velocity. As the lateral force is a major feedback for the driver on the driving style, we assume that it is also a good identifier. However, to avoid identification of the driver via longitudinal driving behavior, we restricted ourselves to only the mean velocity of one window. The feature calculation hence results in a vector $\vec{F}_w = [f_w[0], \dots, f_w[f], \dots, f_w[\hat{h}_{opt}]]^T$ of length $\hat{h}_{opt} + 1$ for each window \vec{W}_w . For each segment, a sequence of feature vectors is obtained and results in one segment matrix per class (i.e. driver) i :

$$\mathbf{B}_i = [\vec{F}_1 \dots \vec{F}_w \dots \vec{F}_F] \\ = \begin{bmatrix} f_1[0] & \dots & f_w[0] & \dots & f_F[0] \\ \vdots & & \vdots & & \vdots \\ f_1[f] & \dots & f_w[f] & \dots & f_F[f] \\ \vdots & & \vdots & & \vdots \\ f_1[\hat{h}_{opt}] & \dots & f_w[\hat{h}_{opt}] & \dots & f_F[\hat{h}_{opt}] \end{bmatrix}, \quad (3)$$

where F is the number of windows within one segment, i.e. $F = S/\hat{h}_{opt}$. The segment matrices were concatenated so that the model was fed by a $[32, F, \hat{h}_{opt}]$ tensor. We chose a batch size of 32 [33], i.e. 32 segment matrices of uniformly randomly picked drivers were used to create one full batch matrix.

2) *Model*: The neural network was built using the *tensorflow* library³. The network processed the features of the windows sequentially and put them into a two-layer bidirectional recurrent neural network (RNN) initialized with GRU cells of size 512. A GRU cell consists of an update gate (z_j) and a reset gate (r_j) that update given input \vec{F}_w and hidden state h_i as

$$z_j = \sigma \left([\mathbf{W}_z \vec{F}_w]_j + [\mathbf{U}_z \vec{h}_{i-1}]_j \right), \quad (4)$$

$$r_j = \sigma \left([\mathbf{W}_r \vec{F}_w]_j + [\mathbf{U}_r \vec{h}_{i-1}]_j \right), \quad (5)$$

Where \mathbf{W} and \mathbf{U} are weight matrices which are learned and σ is the logistic sigmoid function [34]. The hidden state is then updated by:

$$h_{i,j} = z_j h_{i-1,j} + (1 - z_j) \cdot \sigma \left([\mathbf{W} \vec{F}_w]_j + [\mathbf{U} (\vec{r} \odot \vec{h}_{i-1})]_j \right), \quad (6)$$

where \odot denotes the Hadamard-, or element-wise product.

The 512 outputs of the first GRU are fed into the second GRU stage, returning also 512 outputs. It should be emphasized that we do not further process the outputs of GRU. Instead, we use GRU as an encoding method and treat the hidden states of both GRU directions as a summary of a segment, namely the summary of a driver's steering wheel behavior. These hidden states are then reduced by a fully connected (FC) layer to a softmax vector of length of the number of classes. We name the softmax vector the *vote vector*. In addition, to boost the performance of the network, we invoke a voting mechanism in our model, which is inspired by [3] and [6]. Conventionally, in classification tasks a RNN or its variants (GRU, LSTM) only generates one decision at the end of a input sequence, which could be vulnerable due to vanishing and exploding gradient problems and hence lead to a decision that biases towards the most recent inputs. Gradient vanishing and exploding problem can be partially solved by GRU- or LSTM variants of RNNs [35]. The scheme of our neural network is provided in Figure 4.

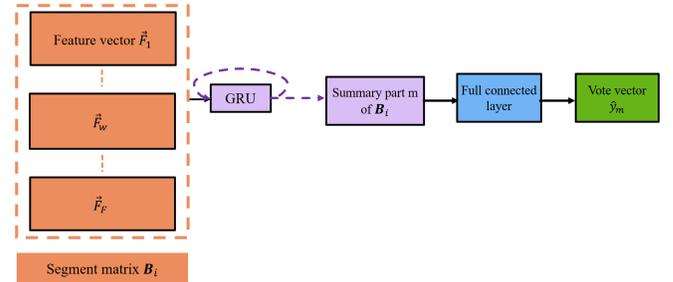


Fig. 4: Scheme of our neural network invoking voting mechanism

As mentioned in the previous paragraph, the hidden states of the GRU, which are reduced to vote vectors, are further processed by a voting mechanism. In the testing and training phase, for every 6 windows, a vote vector is generated from current hidden states of the network. We use only 6 windows to avoid overfitting to the traffic or road topology [29]. These vote vectors are compared with the ground-truth one hot vector \vec{y}_i with cross entropy as the loss function. Hence, F feature vectors in one segment matrix (c.f. Equation (3)) lead to $M = \frac{F}{6}$ vote vectors. The loss of the vote vectors is then aggregated and the parameters of the network are updated by back-propagation. The training was done using the RMSPropOptimizer with a learning rate of 10^{-4} , a state keep probability for the GRUs of 0.7, and a L2-norm weight regularization to avoid overfitting of $\lambda = 10^{-3}$.

³We used the python3 library of tensorflow in it's versions 1.13.1, <https://www.tensorflow.org/>

IV. RESULTS

In this section we present the results that we achieved using the neural network from Section III-C2. First we will give results motivating our approach to design the window size of a neural network via the outfading ACF and the distribution of the lags h_{cor} . Afterwards, the dependency of the model's accuracies and the required observation time are presented. Finally we investigate the confusion matrix of the model's identification.

We required at least 240 minutes of driving data for each driver to train the model and at least 30 minutes of driving data for testing. The testing and training data did not overlap. Through the segment matrix representation, we ensure the same amount of training and testing data for each class.

A. \hat{h}_{opt} - Optimal window size

We limited the number of drivers to the most common class size of 15 and a testing time of 15 minutes. For each window size, at least seven sets of drivers were picked at random from our dataset. Additionally, we used the same set of randomly chosen drivers and trips for each setting to ensure reproducibility. Since the lag h is distributed from 0.5 seconds to 20 seconds, with a mean of ≈ 5.8 seconds, we analyzed the performance of our network with window sizes between 2.5 and 10 seconds. In Figure 5, the accuracies and standard deviations are shown. Additionally to the average accuracy for each window size, the standard deviation is given by the light blue region. It can be seen that the accuracy has its peak

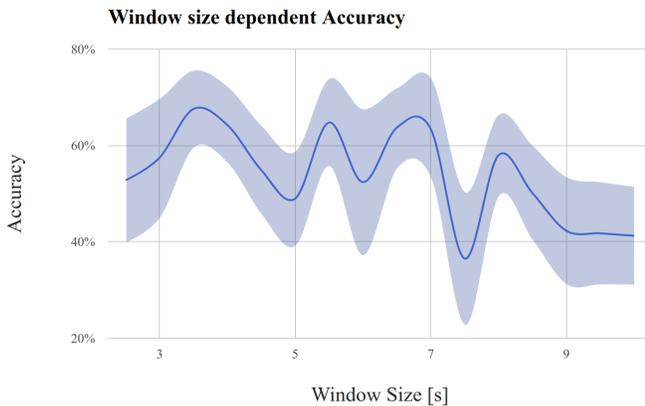


Fig. 5: Accuracy of our neural net for different window sizes for 15 drivers.

around the mode of the histogram of the lag h at 3.5 to 4 seconds with approximately 67%. Although high accuracies of over 60%, on average, were also achieved for 5.5, 6.5 and 7.0 seconds, the best performance was at 3.5 seconds. Above 7 seconds, the accuracies drop significantly to below 60%, hence we limited our analyses to a maximum window size of 10 seconds. Similar significant drops of accuracies were observed for window sizes below 3.5 seconds. Window sizes below 2.5 seconds were not considered as they lead to an increased size of segment matrices and calculations become too complex.

B. Replication of the state of the art

Besides evaluating the newly introduced neural net, we further replicated the work presented in [3]. The remaining major differences are a lower resampling frequency of 10 Hz due to a lower sampling rate at the CAN-bus and our naturalistic driving data set compared to the controlled driving data set. The strong claims of 83% accuracy after a training phase of 90 minutes of driving data per user [3] could not be replicated in our setting. Figure 6 shows the accuracies of the replicated studies using 240 minutes of driving data per class. It can be seen that the baseline of a random guess at

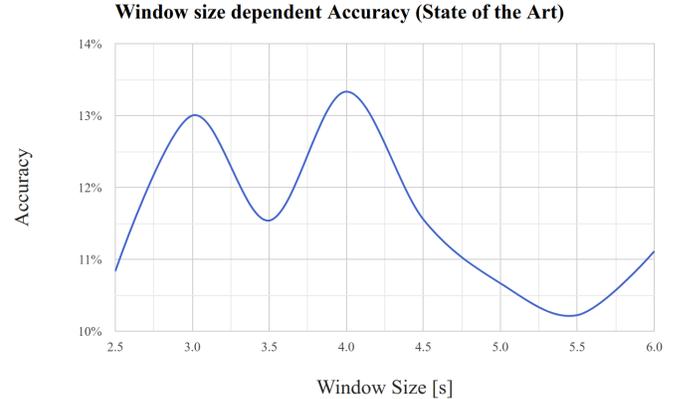


Fig. 6: Accuracy of the replicated state of the art models for different window sizes for 15 drivers.

7% was improved, and the best performance was achieved with a window size of approximately 4 seconds similar to results presented in Figure 5. However, with an accuracy of less than 15%, the difference to our introduced model is significantly lower. This is in line with previous replicated research. In [6], a difference of over 55% between naturalistic- and controlled driving data is presented. A similar difference of approximately 50% can be seen comparing Figure 5 and 6.

C. Model performance dependency on observation time

In the following we use the optimal window size of 3.5 seconds found in the previous results. Further we used a training set size of approximately 240 minutes per driver following [9]. We investigate the dependency of the identification accuracy and the observed driving time. The observation time is limited to 15 minutes, since a longer observation is impractical. In Figure 7 the identification accuracy is shown for a different number of drivers ranging from 15 to the full dataset of 72 drivers. As expected we see that the accuracy increases with a smaller set size of drivers. Further, a clear dependency of observation time and accuracy can be observed. The longer the neural network “observes” a driver, the higher the identification accuracy. Further we can observe that the accuracy tends to stabilize after 12 minutes for 15 and 30 drivers, but not for 72 drivers, indicating that a larger evaluation set size is required for a larger class size to reach a final vote. After the first vote vector (21 seconds), accuracies lie at approximately 25% for 15 drivers, 10% for 30% for 30 drivers, and 8% for

72 drivers. These accuracies increase to 76% for 15 drivers, increasing the state of the art by a factor of 5 and random guess by a factor of 11 after 42 vote vectors, or almost 15 minutes. For larger class sizes the accuracies drop to 41% (30 drivers) and 36% (72 drivers).

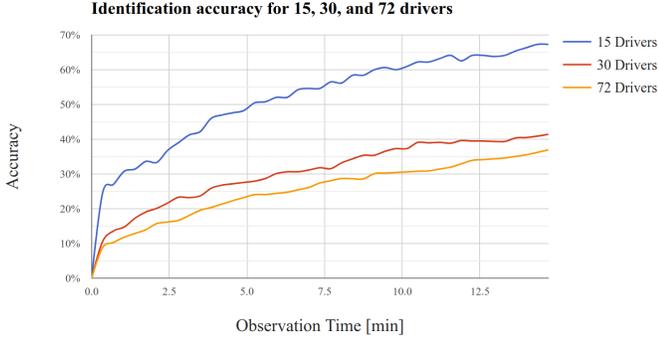


Fig. 7: Accuracy of our neural net for 15, 30, and 72 drivers over observation time.

D. Confusion Matrix

Figure 8 shows the normalized confusion matrix of 15 drivers after training the algorithm with 240 minutes of driving data averaged over 315 different 15-minute segments. (i.e. 21 segments per driver) We see that three of the 15 drivers were identified correctly for each of the 21 segments. It can also be seen that for 12 of the 15 drivers, the diagonal element is the strongest element, indicating a major correct identification. The remaining three drivers (Driver 5, 8, and 12) hence are on average mostly misidentified. As described in Section III, the drivers were stationed in different regions with a strongly varying road topology. Only Driver 12 was extremely misidentified towards Driver 9. However, the two drivers were stationed at different locations, therefore similar steering behavior is more likely than misfitting of the model towards the region. Finally we can observe that the neural network was strongly misfitted towards Driver 10, creating a “sink.” Again, none of the misidentified drivers were stationed in the same region as Driver 10, therefore a similar driving style is most likely the reason for this over-fitting. We neglected the confusion matrix for 72 drivers in this section for readability reasons. The same effects (with lower accuracies, c.f. Figure 7) can be observed.

V. DISCUSSION AND OUTLOOK

Research into the growing field of automated vehicle person identification is motivated by various use-cases, including privacy, personalized service models, and the potential to enable advanced applications such as health and well-being related products and services. The results presented in this work contribute to the understanding of the current status of driver identification systems.

The results should be seen in the light of the study limitations. Despite the unique size of the dataset, we acknowledge that the results are based on a naturalistic field test of professional road assistance drivers. As such, generalization of the

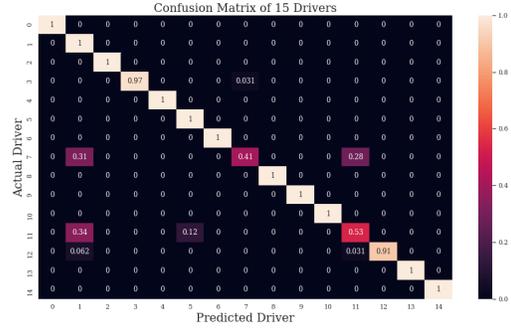


Fig. 8: Confusion matrix of our neural network for 15 drivers.

results remains a core challenge that should be addressed in future work. However, we expect that there is reduced variance in the feature distribution of professional drivers when compared to a sample of more ordinary drivers. As the homogeneity of the sample may have limited the prediction results, greater identification accuracy might be possible in a more general setting. Thus, we encourage other researchers to validate the strong claims made in this work on more generalizable data sets. Further, as we limited ourselves to an extremely low dimension of features, only using the logarithmic Fourier-transformation of the steering wheel and the average velocity-per-window, more precise and faster identification could be expected with greater feature engineering.

In summary, this work is based on an extensive literature review that, to our knowledge, has not been previously conducted by the field. We identified driver identification based on naturalistic driving data using only steering wheel signal as a research gap. Large discrepancies in literature on the choice of the window size motivated our signal-processing-based approach of an optimal hyper parameter choice. The newly introduced approach is a lightweight method, which could find further applications outside this field in time-series-dependent tasks. Additionally, we present a GRU-based neural network identification setup that improves the results from state-of-the-art methods by a factor of almost 5. The analysis goes beyond the traditional sample size of previous studies, and considers the largest fleet-based setting of 72 drivers under naturalistic conditions. Finally, to consolidate our contribution, we investigated the confusion matrix and confirm that the model did not have bias toward the region of driving. Rather, drivers were misidentified because of their similar driving style, suggesting that a more heterogeneous driver set would lead to higher accuracies.

Based on the improvements achieved, we conclude that single-sensor-tailored approaches can increase accuracies significantly, while at the same time avoid misidentification bias from the car or the region. Hence, in future work, this approach should be combined with existing methods, and for additional sensors, in order to leverage the maximum results and demonstrate the optimal real-world capabilities of data enabled driver identification.

Author	Test Environment	ML Method	Window Size	Algorithm Input	Training Set	Class Size	Accuracy
Gahr et al. 2018 [6]	Naturalistic driving	Random-forest	Event based	CAN-bus Brake pedal signal	120 - 800 brakings	5	95.0% - 99.5%
					≈ 50min - 5h30min per driver	15	80.0% - 92.0%
						50	70.0% - 85.0%
Hallac et al. 2018 [5]	Naturalistic driving	Gated recurrent units	1 Second	665 CAN-bus signals	≈ 23 hours per driver	56	51.3%
Hallac et al. 2016 [4]	Naturalistic driving	Random-forest	Event based	12 CAN-bus signals	11-34 turnings per driver	2	76.9%
						3	59.4%
						4	55.2%
						5	50.1%
Wang et al. 2017 [9]	Naturalistic driving	Random-forest	5 Seconds	CAN-bus Signals: Speed,Torque, Brake Pressure, Throttel Position, Steering Angle	4 hours per driver	30	93% - 100%
Dong et al. 2017 [29]	Naturalistic driving	ARNet	4 Seconds	6 GPS-speed derived Signals	160 trips	50	40%
Zhang et al. 2016 [28]	Naturalistic driving	SVM	30 Seconds	CAN-bus & Smartphone Data	≈1.25 hours	14	29%
Wakita et al. 2005 [10]	Simulator study	Gaussian Mixed Model (GMM)	0.6 Seconds	CAN-bus signals & following distance	10 min	12	81%
	Controlled driving study					30	73%
Miyajima et al. 2006 [14], [15]	Simulator study	GMM	0.8 Seconds	CAN-bus signals & following distance	3 min	276	89.6%
	Controlled driving study						76.8%
Del Campo et al. 2014 [16]	Controlled driving study [17]	Multi-layer Perceptron	2 Seconds	Gas- & Brake pedal CAN-bus signal	≈ 8 min	3 - 5	84% - 70%
Martinez et al. 2016 [18]	Controlled driving study [17]	Extreme Learning Machine	128 Seconds	14 CAN-bus signals	N.A.	2 - 11	97% - 85%
Enev et al. 2016 [3]	Controlled driving study	Random-forest	3.5 Seconds	15+ CAN-bus signals	90 min	15	100%
					15 min		100%
				Brake Pedal Signal	90 min		100%
					15 min		87.33%
				Steering Wheel Signal	90 min		83.33%
					15 min		64.67%
Burton et al. 2016 [19]	Simulator study	Decision Trees, SVM, kNN	10 Seconds	Vehicle coordinates & CAN-bus signals	16 min	10	100%
Jafarnejad et al. 2017 [23]	Controlled driving study	Random-forest w/ variations & SVM	15 Seconds	CAN-bus Signals: Speed, Torque, Throttel, Yaw Rate, Steering Angle	≈ 27.5 min	5	95%
						15	89%
						35	82%
Ezzini et al. 2018 [8]	Controlled driving study [20]	Extra Trees	N.A. / Variable	51 CAN-bus signals	5 min	10	99.92%
	Controlled driving study [21]			GPS, brightness, acc., & physiological signals	1 min	6	99.99%
	Controlled driving study [22]				2 min - 5min		100.0%
Jeong et al. 2018 [24]	Controlled driving study	CNN, SVM, ELM	0.25 Seconds	CAN-bus	75s per driver	4	88%
Marchegiani et al. 2018 [7]	Controlled driving study	SVM & GMM	5 & 15 Seconds	CAN-bus: Throttle & Brake	≈190 min	4	83%
Rettore et al. 2018 [25]	Controlled-, and driving study	Extra Trees	120 Seconds	19 CAN-bus, Smartphone, & Virtual sensor signals	≈2.8h per driver	10	98%
Bernardi et al. 2018 [26]	Controlled driving study	Multi-layer Perceptron	1 Minute	16 CAN-bus & Smartphone Signals	N.A.	10	95%
Luo et al. 2018 [27]	Controlled driving study	Random Forest	N.A.	300 CAN-bus signals	≈120min	4	89%

TABLE I: Detailed overview of the related work.

REFERENCES

- [1] S. M. Iranmanesh, H. N. Mahjoub, H. Kazemi, and Y. P. Fallah, "An adaptive forward collision warning framework design based on driver distraction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 12, pp. 3925–3934, Dec 2018.
- [2] L. F. Wurster, J. M. Correia, A. Woodyer, and C. Eschinger, "Market Trends: Prepare Now to Smooth the Inevitable Transition to a Subscription-Based Business Model." Gartner, Tech. Rep., 2015.
- [3] M. Enev, a. Takakuwa, K. Koscher, and T. Kohno, "Automobile Driver Fingerprinting," *Proceedings on Privacy Enhancing Technologies*, vol. 2016, no. 1, pp. 34–50, 2016.
- [4] 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 *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 953–958.
- [5] D. Hallac, S. Bhooshan, M. Chen, K. Abida, R. Sosič, and J. Leskovec, "Drive2vec: Multiscale state-space embedding of vehicular sensor data," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, Nov 2018, pp. 3233–3238.
- [6] B. Gahr, B. Ryder, A. Dahlinger, and F. Wortmann, "Driver identification via brake pedal signals — a replication and advancement of existing techniques," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, Nov 2018, pp. 1415–1420.
- [7] L. Marchegiani and I. Posner, "Long-term driving behaviour modelling for driver identification," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, Nov 2018, pp. 913–919.
- [8] S. Ezzini, I. Berrada, and M. Ghogho, "Who is behind the wheel? driver identification and fingerprinting," *Journal of Big Data*, vol. 5, no. 1, p. 9, Feb 2018. [Online]. Available: <https://doi.org/10.1186/s40537-018-0118-7>
- [9] B. Wang, S. Panigrahi, M. Narsude, and A. Mohanty, "Driver identification using vehicle telematics data," in *SAE Technical Paper*. SAE International, 03 2017. [Online]. Available: <https://doi.org/10.4271/2017-01-1372>
- [10] T. Wakita, K. Ozawa, C. Miyajima, K. Igarashi, K. Itou, K. Takeda, and F. Itakura, "Driver identification using driving behavior signals," in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, vol. 2005, 2005, pp. 907–912.
- [11] A. Riener and A. Ferscha, "Supporting Implicit Human-to-Vehicle Interaction : Driver Identification from Sitting Postures," in *Proceedings of the First Annual International Symposium on Vehicular Computing Systems. July 22-24, 2008, Dublin*, 2008, pp. 1–10.
- [12] J.-D. Wu and S. H. Ye, "Driver identification using finger-vein patterns with Radon transform and neural network," *Expert Systems with Applications*, vol. 36, no. 3 PART 2, pp. 5793–5799, 2009.
- [13] J.-D. Wu and S.-H. Ye, "Driver identification based on voice signal using continuous wavelet transform and artificial neural network techniques," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1061–1069, 2009.
- [14] C. Miyajima, Y. Nishiwaki, K. Ozawa, T. Wakita, K. Itou, and K. Takeda, "Cepstral Analysis of Driving Behavioral Signals for Driver Identification," *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 5, no. October 2015, pp. 6–9, 2006.
- [15] C. Miyajima, Y. Nishiwaki, K. Ozawa, T. Wakita, K. Itou, K. Takeda, and F. Itakura, "Driver modeling based on driving behavior and its evaluation in driver identification," *Proceedings of the IEEE*, vol. 95, no. 2, pp. 427–437, 2007.
- [16] I. Del Campo, R. Finker, M. Martinez, J. Echanobe, and F. Doctor, "A real-time driver identification system based on artificial neural networks and cepstral analysis," in *International Joint Conference on Neural Networks, IJCNN 2014*. IEEE, 9 2014, pp. 1848–1855.
- [17] H. Abut, H. Erdoğan, A. Erçil, A. B. Çürüklü, H. C. Koman, F. Tas, A. Ö. Argunşah, B. Akan, H. Karabalkan, E. Çökelek *et al.*, "Data collection with "uyanik": too much pain; but gains are coming," 2007.
- [18] M. V. Martínez, J. Echanobe, and I. del Campo, "Driver identification and impostor detection based on driving behavior signals," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 372–378.
- [19] A. Burton, T. Parikh, S. Mascarenhas, J. Zhang, J. Voris, N. S. Artan, and W. Li, "Driver identification and authentication with active behavior modeling," in *2016 12th International Conference on Network and Service Management (CNSM)*, Oct 2016, pp. 388–393.
- [20] B. I. Kwak, J. Woo, and H. K. Kim, "Know your master: Driver profiling-based anti-theft method," in *PST 2016*, 2016.
- [21] S. Schneegass, B. Pflöging, N. Broy, F. Heinrich, and A. Schmidt, "A data set of real world driving to assess driver workload," in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ser. *AutomotiveUI '13*. New York, NY, USA: ACM, 2013, pp. 150–157. [Online]. Available: <http://doi.acm.org/10.1145/2516540.2516561>
- [22] E. Romera, L. M. Bergasa, and R. Arroyo, "Need data for driver behaviour analysis? presenting the public uah-driveset," in *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*. IEEE, 2016, pp. 387–392.
- [23] S. Jafarnejad, G. Castignani, and T. Engel, "Towards a real-time driver identification mechanism based on driving sensing data," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Oct 2017, pp. 1–7.
- [24] D. Jeong, M. Kim, K. Kim, T. Kim, J. Jin, C. Lee, and S. Lim, "Real-time driver identification using vehicular big data and deep learning," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, Nov 2018, pp. 123–130.
- [25] P. H. L. Rettore, A. B. Campolina, A. Souza, G. Maia, L. A. Villas, and A. A. F. Loureiro, "Driver authentication in vanets based on intra-vehicular sensor data," in *2018 IEEE Symposium on Computers and Communications (ISCC)*, June 2018, pp. 00078–00083.
- [26] M. L. Bernardi, M. Cimitile, F. Martinelli, and F. Mercaldo, "Driver identification: a time series classification approach," in *2018 International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1–7.
- [27] D. Luo, J. Lu, and G. Guo, "Driver identification using multivariate in-vehicle time series data," in *WCX World Congress Experience*. SAE International, apr 2018. [Online]. Available: <https://doi.org/10.4271/2018-01-1198>
- [28] C. Zhang, M. Patel, S. Buthpitiya, K. Lyons, B. Harrison, and G. D. Abowd, "Driver classification based on driving behaviors," in *Proceedings of the 21st International Conference on Intelligent User Interfaces*, ser. *IUI '16*. New York, NY, USA: ACM, 2016, pp. 80–84. [Online]. Available: <http://doi.acm.org/10.1145/2856767.2856806>
- [29] W. Dong, T. Yuan, K. Yang, C. Li, and S. Zhang, "Autoencoder regularized network for driving style representation learning," *CoRR*, vol. abs/1701.01272, 2017. [Online]. Available: <http://arxiv.org/abs/1701.01272>
- [30] W. Fuller, J. Watson, and W. Fuller, *Introduction to Statistical Time Series*, ser. *Wiley Series in Probability and Statistics*. Wiley, 1996. [Online]. Available: <https://books.google.ch/books?id=wyRhjmAPQIYC>
- [31] M. Zihlmann, D. Perekrestenko, and M. Tschannen, "Convolutional recurrent neural networks for electrocardiogram classification," in *Computing in Cardiology (CinC)*, Sep. 2017. [Online]. Available: <http://www.nari.ee.ethz.ch/comth/pubs/p/cinc2017>
- [32] P. Koehn and M. Eckrich, "Active steering - the bmw approach towards modern steering technology," in *SAE 2004 World Congress & Exhibition*. SAE International, mar 2004. [Online]. Available: <https://doi.org/10.4271/2004-01-1105>
- [33] D. Masters and C. Luschi, "Revisiting small batch training for deep neural networks," *CoRR*, vol. abs/1804.07612, 2018. [Online]. Available: <http://arxiv.org/abs/1804.07612>
- [34] K. Cho, B. van Merriënboer, Ç. Gülçehre, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *CoRR*, vol. abs/1406.1078, 2014. [Online]. Available: <http://arxiv.org/abs/1406.1078>
- [35] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.



Bernhard Gahr Bernhard Gahr is a PhD candidate at the Bosch IoT Lab at the University of St.Gallen (HSG) and investigates the potential of the connected car in the field of personalization. He holds a diploma (M.Sc.) in electrical engineering and information technology from ETH Zurich. In addition, he completed an international stay for one year, studying at the LTH in Lund, Sweden. Before joining the Bosch IoT lab, he worked for Open Systems on network security and for Bosch on combustion engine fuel injection systems.



André Dahlinger André Dahlinger received his doctoral degree from the University of St.Gallen (HSG) in 2018 and investigated the potential of the connected car for improving driving performance and car security. He holds a Diploma (M.Sc.) in psychology and a B.Sc. in economics from the Philipps-University of Marburg. After spending one semester studying economics in Salamanca, Spain, he wrote his Diploma thesis in the field of behavioral economics (The Effect of Ownership on Decisions under Risk) and his Bachelor thesis in microeconomics (Modeling Emotions in Microeconomics).



Shu Liu Shu Liu is a PhD candidate at the Bosch IoT Lab at ETH Zurich. His research interest lies in the area of Machine Learning and Computer Vision in the context of Internet of Things / V2V and Autonomous Driving. He holds a M.Sc in Electrical Engineering from ETH Zurich and a B.Sc from TU Munich. Before joining the Bosch IoT Lab he worked as a Research Assistant at the Computer Vision and Geometry Group of ETH Zurich in cooperation with MIT.



Benjamin Ryder Benjamin Ryder received his doctoral degree from ETH Zurich in 2018 and investigated the potential of the connected car for improving driving performance and car security. Ben graduated with an MEng in Computer Science from Imperial College London, United Kingdom, where his masters thesis investigated the potential of the Microsoft Kinect as a touch screen input device. Benjamin undertook a research placement at the Hamlyn Centre for Robotic Surgery, and after graduating worked for two years as a Technical Consultant with BAE Systems on a variety of projects focusing on autonomously detecting fraud for Insurance and Finance companies.



Kevin Koch Kevin Koch is a PhD candidate at the Bosch IoT Lab at the University of St.Gallen (HSG). His research focuses on the potential of the connected car for driver state analysis and driving performance. Kevin holds a Diploma (M.Sc.) in Information Systems from Technische Universität München (TUM) and a bachelor degree (B.Sc.) in Business Information Technology from Frankfurt School of Finance & Management. In addition, he completed several international stays, including a research stay at the National Institute of Informatics

in Tokyo (NII).



Elgar Fleisch Elgar Fleisch, studied information systems at the University of Vienna (Austria). He received his doctoral degree from the University of Economics and Business Administration Vienna in 1993 for his work in artificial intelligence in production planning. Between 1994 and 2000 he worked as a postdoc in information management at the University of St. Gallen (Switzerland). In 2000 Elgar Fleisch was elected Associate Professor at the University of St. Gallen, becoming a full professor in 2002. In 2004 he was elected Professor for Information Management at ETH Zurich and started a double professorship with the University of St. Gallen.



Filipe Barata Filipe Barata is a Ph.D. candidate and doctoral researcher at the Center for Digital Health Interventions at ETH Zurich. He employs artificial intelligence in the emerging field of patient-centered health information systems. In particular, he is investigating how to develop an effective, smartphone-based self-monitoring system for asthma. He holds a diploma (M.Sc.) in electrical engineering and information technology from ETH Zurich. Before joining the Center for Health Interventions, he worked in the research department of Bruker Biospin, where he developed algorithms for the autonomous interpretation of NMR spectra. In his professional career he gained experience in several research laboratories, including the Department of Biomedical Data Science at the Geisel School of Medicine at Dartmouth College and the R&D department of LMS International (Siemens).



Felix Wortmann Felix Wortmann is Assistant Professor of Technology Management at the University of St. Gallen. Furthermore, he is the Scientific Director of the Bosch IoT Lab, a research collaboration between Bosch, the University of St. Gallen and ETH Zurich. His research interests include the Internet of Things, blockchain, big data, and digital innovations in mobility, energy, and health. Felix Wortmann received a BScIS and MScIS from the University of Muenster, Germany, and a PhD in Management from the University of St. Gallen. He gained several years of industry experience in a German-based multinational software corporation.