Driving Anomaly Detection with Conditional Generative Adversarial Network using Physiological and CAN-Bus Data

Yuning Qiu
University of Texas at Dallas
Richardson, Texas
yxq180000@utdallas.edu

Teruhisa Misu
Honda Research Institute
Mountain View, California
tmisu@honda-ri.com

Carlos Busso
University of Texas at Dallas
Richardson, Texas
busso@utdallas.edu

Figure 1: Examples of hazard driving scenarios that can trigger physiological reactions or unexpected maneuvers from the drivers. The proposed unsupervised anomaly detection system aims to identify these events.

ABSTRACT

New developments in *advanced driver assistance systems* (ADAS) can help drivers deal with risky driving maneuvers, preventing potential hazard scenarios. A key challenge in these systems is to determine when to intervene. While there are situations where the needs for intervention or feedback is clear (e.g., lane departure), it is often difficult to determine scenarios that deviate from normal driving conditions. These scenarios can appear due to errors by the drivers, presence of pedestrian or bicycles, or maneuvers from other vehicles. We formulate this problem as a driving anomaly detection, where the goal is to automatically identify cases that require intervention. Towards addressing this challenging but important goal, we propose a multimodal system that considers (1) physiological signals from the driver, and (2) vehicle information obtained from the *controller area network* (CAN) bus sensor. The system relies on conditional *generative adversarial networks* (GAN) where the models are constrained by the signals previously observed. The difference of the scores in the discriminator between the predicted and actual signals is used as a metric for detecting driving anomalies. We collected and annotated a novel dataset for driving anomaly detection tasks, which is used to validate our proposed models. We present the analysis of the results, and perceptual evaluations which demonstrate the discriminative power of this unsupervised approach for detecting driving anomalies.

CCS CONCEPTS

• **Computing methodologies → Scene anomaly detection**; **Anomaly detection**; *Neural networks.*

KEYWORDS

ADAS, anomaly detection, conditional GAN, physiological data

ACM Reference Format:
Yuning Qiu, Teruhisa Misu, and Carlos Busso. 2019. Driving Anomaly Detection with Conditional Generative Adversarial Network using Physiological and CAN-Bus Data. In *2019 International Conference on Multimodal Interaction (ICMI '19), October 14–18, 2019, Suzhou, China.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3340555.3353749

1 INTRODUCTION

Over the past decade, vehicles equipped with *advanced driver assistance systems* (ADAS) have made important safety improvements,

ICMI '19, October 14–18, 2019, Suzhou, China

ACM ISBN 978-1-4503-6860-5/19/05…$15.00
https://doi.org/10.1145/3340555.3353749

enhancing the driving experience on the roads. Examples include *forward collision warning* (FCW), *intelligent speed advice* (ISA), collision avoidance system, and blind spot monitor. Studies have also attempted to directly detect driver distractions [5, 10, 13], or detect primary tasks need to safely complete a driving maneuver (e.g., mirror-checking actions [11]) Most ADAS functions get activated when drivers fail to properly drive or react to changes in the driving environment. A key step in these systems is to understand when to intervene. New advances in ADAS will require to identify driving anomalies that require intervention, such as the examples shown in Figure 1.

In this work, we define driving anomalies as events that deviate from expected driver behaviors that can lead to hazard situations. Examples of a driver-related anomalies includes abrupt changes on driving maneuvers, missing primary tasks required to complete a driving maneuver (e.g., checking mirrors before turning [11, 12]), and lack of awareness of the presence of objects, pedestrians, or other vehicles. Examples of scene-related anomalies include hazard actions from other vehicles, and unexpected changes on the road that leads to hazard scenarios (e.g., constructions on the road). Our aim is to automatically estimate driver anomaly using multimodal sensors with unsupervised methods. In particular, this study considers data from the driver (e.g., physiological signals), and from the vehicle via the *controller area network* (CAN) bus (e.g., acceleration, steering wheel position), leveraging the extensive naturalistic driving data introduced in this study. This approach is appealing since it not only does not require expensive and time-consuming annotations, but also can inform of non-intuitive types of driving anomalies that cannot be easily tabulated with pre-defined rules.

The key idea behind the proposed approach is to automatically identify driving segments that deviate from normal or expected patterns. Our formulation creates predictions conditioned on data from previous segments. These predictions are compared with the actual data, quantifying the deviations. We implement these ideas with conditional *generative adversarial networks* (GANs). Our implementation of conditional GAN creates predictions of the data for the following six seconds, conditioned on the previous six seconds. The discriminator and the generator are conditioned on the data preceding the target window. We create a driving anomaly score by leveraging the output layer of the discriminator. We present to the discriminator the predicted samples creates by the generator and the actual samples observed during the target six seconds. We estimate the difference of the sigmoid outputs for both signals, using this score as our driving anomaly metric. This metric increases its value when the difference between the predicted and actual data increases, indicating that something unexpected happened. This unsupervised approach learns useful data representation without the need of labeled data, identifying segments that deviate from normal driving scenarios.

To evaluate the proposed approach, we collected 250 hours of naturalistic urban driving data, which we refer to as *driving anomaly dataset* (DAD). This work uses a subset of 48 hours. We consider three analyses to assess the performance of the proposed driver anomaly detection approach. The first evaluation analyzes the difference between the distributions of our anomaly score for two subsets of the data: the *candidate* and *normal* sets. The *candidate* set includes segments annotated with avoid on-road pedestrian, avoid on-road bicyclist, avoid parked vehicle, and traffic rule violation. The *normal* set are 400 randomly selected segments without any annotation or driving maneuver. Our analysis reveals that the anomaly scores for *candidate* segments are generally higher than the anomaly scores for *normal* segments. The second evaluation compares the 100 segments with the highest anomaly scores with 100 randomly selected segments. We compare the annotations provided in the dataset overlapping with these two sets. We observe more risky events annotated over segments with high anomaly scores. The third evaluation validates this result with perceptual evaluations. We select the 40 segments with the highest anomaly scores, and 40 segments randomly selected. We ask four evaluators to watch these videos, judging their risk and familiarity levels. The result reveals that videos with the highest anomaly scores are perceived as more risky and less common than randomly selected videos. Collectively, these results indicate that the proposed unsupervised anomaly scores using conditional GAN are effective in detecting driving recordings that deviate from normal recordings.

2 RELATED WORK

2.1 Driving Anomaly Detection

Driving anomaly detection is an important problem. Studies have proposed anomaly detection approaches in very specific problems by setting thresholds. Malta et al.[16] proposed an anomaly detection model based on brake pedal operations and thresholds on the vehicle speed. They consider a hazardous scenario when the pressure on the brake pedal was high while the mean velocity was above a given threshold. Zhao et al.[24] detected aggressive driving events using steering wheel information and acceleration information collected from smartphones. They set acceleration thresholds that depended on the steering wheel angles. The thresholds were more sensitive when the steering wheel angle was high. These models based on predefined rules can only work well for simple cases. When the driving environment is complex, however, they can overestimate the risk even when the drivers are properly controlling the vehicle (e.g., changes in acceleration associated with overtaking another vehicle), or underestimate risk when the thresholds are not satisfied.

An alternative approach is to detect anomalies using machine-learning algorithms. Selmanaj et al.[22] proposed an anomaly detection model based on two classifiers. As input, they used vertical and horizontal accelerations of a motorcycle collected during naturalistic driving recordings. The data are labeled as normal, irregular, and hazard according to the road conditions (baseline road without any specific feature, irregular roads with holes or speed bumps, hazard roads with rough bumps or sharp turns). The first classifier discriminates the samples as normal or anomaly. The second classifier classifies the anomaly samples into two sets, irregular and hazard, corresponding to the driving conditions. The results were validated with simulated data. Chen et al.[1] proposed a SVM-based approach to classify six different abnormal behaviors: weaving, swerving, sideslipping, fast U-turn, turning with a wide radius and sudden brakes. The study used the accelerometer and orientation sensors of a smartphone. However, due to the complexity and diversity of driving anomalies, designing an exhaustive classification-based solution that can precisely identify all kinds of driving anomaly scenarios can be quite difficult and time-consuming.

Clustering based approaches have also been used for driving anomaly detection. Zheng et al.[25] proposed an unsupervised clustering method using accelerator sensor data of a smartphone. They considered the outliers on the clustering map as the anomalies. However, it is difficult to identify meaningful clusters as the dimension of the feature space increases.

2.2 Physiological Signals & Driving Maneuvers

Studies have shown the close connection between human's physiological signals and the autonomous nervous system [6, 23]. Changes in our mental state, caused by stress, rage, or unexpected events, are reflected in changes in our physiological signals. Haruyuki [6] used the ratio between low and high frequency components of the heart rate power spectrum to measure a person's stress level. The study of Timmons et al.[23] used data extracted from the *eletrocardiogram* (ECG) and *electrodermal activity* (EDA) signals to detect changes in arousal to monitor conflict between individuals.

Studies have analyzed the relationship between the drivers' mental state and driving maneuvers with physiological sensors. The studies of Healey and Picard [7] and Nishigaki et al.[18] revealed that driving can increase the drivers' cognitive workload and mental stress level. Following this direction, many studies have evaluated the relationship between the driver's physiological data and driving maneuvers. Li et al. [14] used *heart rate* (HR) and *breath rate* (BR) signals to connect the driver's mental state with driving events. Murphey et al. [15] designed a lane change detection model based on physiological data.

These studies suggest that changes in physiological signals can be effective cues to identify unexpected driving scenarios that increase the driver's stress levels. We expect to identify driver and vehicle related anomalies by jointly modeling physiological signals with data coming from the vehicle (e.g., CAN-Bus).

2.3 Generative Adversarial Networks

Our proposed models are based on a conditional GAN. GANs are generative models that can create synthetic data that approximate real samples [4]. This goal is achieved by playing an adversarial game between a generative model G and a discriminative model D. During the training procedure, D is trained to discriminate whether a sample comes from real data or from samples generated by G. G is trained to generate plausible samples from noise to confuse D. This adversarial game leads the generator to learn the distribution of the samples that are being generated. An interesting extension of this framework is conditional GANs [2, 8], where the models are also constrained by a given variable that is provided as an input. By adding some conditions as inputs, G can generate fake samples with specific conditions or characteristics rather than a generic sample from a noise distribution.

Li et al.[9] applied GANs on anomaly detection in signal domain for system security issues. They trained a GANs-based model to learn the distribution of signals from sensors and actuators with a Cyber-Physical System working under normal conditions. They used the model to differentiate between normal and abnormal signals, which indicate potential attacked situations.

Some successful examples of conditional GANs applied to different problems include the work of Dai et al.[2], Sadoughi and Busso [21], and Isola et al.[8]. Dai et al.[2] leveraged conditional GAN to generate diverse captions from images. The models were conditioned on the type of image (e.g., drawings versus real pictures). Sadoughi and Busso [21] created speech-driven models for lip movement generation using conditional GAN, which were constrained by the acoustic features. Isola et al.[8] used conditional GAN to generate images with a given style. The condition in this implementation was a picture with the target style. To the best of our knowledge, our paper is the first study that uses conditional GAN as an unsupervised solution for driving anomaly detection.

3 DRIVING ANOMALY DATASET

This study introduces the *driving anomaly dataset* (DAD). We obtained 250 hours of naturalistic driving data. Four drivers collected the data using a Honda Accord. During the data collection, we recorded the driving situation by placing cameras looking at the road and the driver. Those videos are for annotation purposes and are not used for anomaly detection in this work. We also recorded data from the vehicle *controller area network* (CAN)-bus that provides various signals from the vehicle including throttle angle, brake pressure, steering angle, yaw rate and speed at 100 Hz. The physiological signals from the drivers were recorded with wearable devices (Zephyer BioHarness 3 chestband, and Empatica E4). These sensors provide ECG (250 Hz), respiration wave (25 Hz) and skin conductivity (4 Hz) signals. We extract the drivers' *heart rate* (HR), *breath rate* (BR), and *electrodermal activity* (EDA) from these physiological signals. All the sensor signals are synchronized at 30 Hz for convenience. In this paper, we analyze 46 sessions of naturalistic urban driving recordings with a total duration of around 48 hours, which are the sessions that have been currently annotated. We split the 48 hours of driving data into two partitions, using 42 hours for training, and six hours for testing the proposed unsupervised framework.

The DAD is a new completed multimodal dataset and it is used for research purpose for the first time. Unlike others driving dataset such as the KITTI corpus [3] or the HDD corpus [17, 20], our dataset explicitly annotates traffic rule-violations for anomaly detection research purposes. The data was collected in a city in Asia by a local company. By looking into the videos, we find that the traffic conditions in the DAD are much more complicated than the traffic conditions observed in the KITTI corpus (Europe) or the HDD corpus (San Francisco Bay Area, USA) (e.g., more wild drivers and more pedestrians and bicyclists who ignore the traffic rules). These above-mentioned advantages of the DAD motivate us to apply our driving anomaly detection models on this novel corpus.

We manually annotated the corpus with relevant driving events using the software ELAN (Fig. 2). We group some of these annotations to evaluate the models. Table 1 lists the sets. The first set corresponds to the *candidate* set, which includes annotations with hazard scenarios, and traffic violations. We expect that these segments should be more anomalous than other driving segments. The second set is the *maneuver* set, which includes all the annotations related to driving maneuvers. The third set is the *normal* set, which

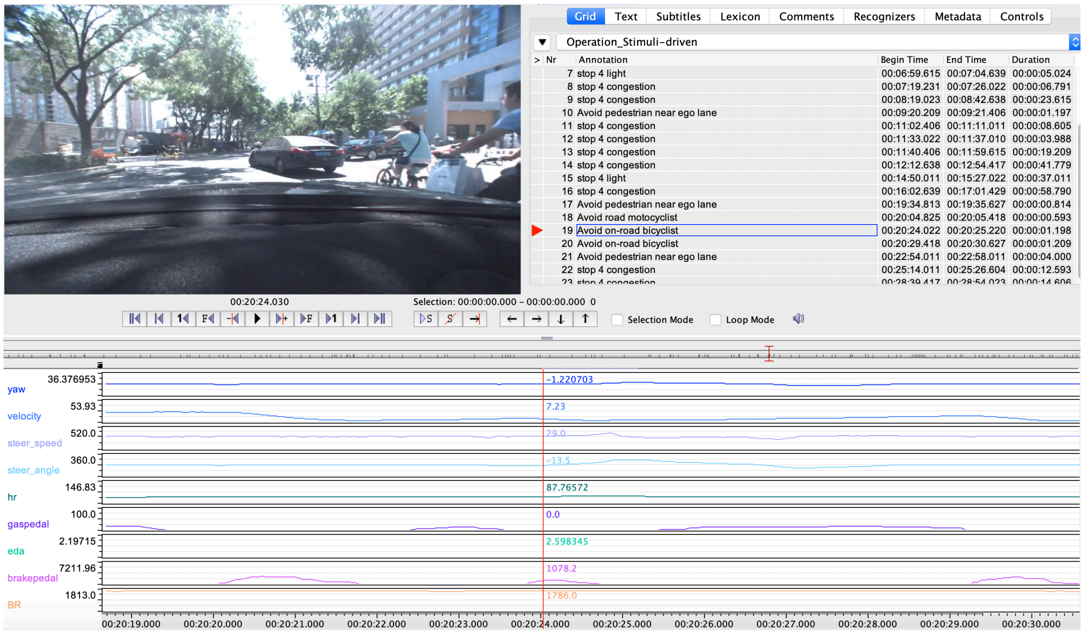


Figure 2: Data interface with the annotations of the recordings using the software ELAN. The physiological signals, CAN-bus data, and annotations are synchronized with the video recording, which shows the road view from the car.

Table 1: Sets considered in this study to evaluate the proposed anomaly detection models. The candidate set is expected to be more anomalous than the other sets.

Sets	Annotations
Candidate	Avoid on-road pedestrian; Avoid pedestrian near ego-lane; Avoid on-road bicyclist; Avoid bicyclist near ego-lane; Avoid on-road motorcyclist; Avoid parked vehicle; traffic rule violation
Maneuver	Intersection passing; Left turn; Right turn; Left lane change; Right lane change; Crosswalk passing; U-turn; Left lane branch; Right lane branch; Merge; Stop for red light; Stop for congestion
Normal	No annotations during the segments

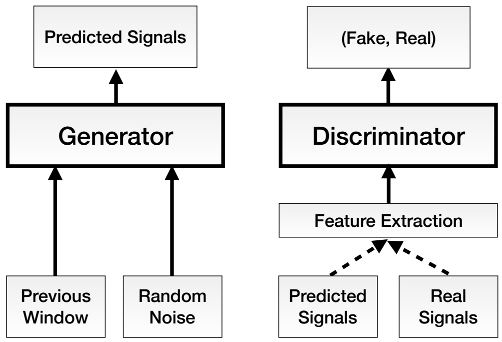
includes driving recordings without any annotations. We use this set as a baseline describing normal driving conditions.

4 PROPOSED CONDITIONAL GAN

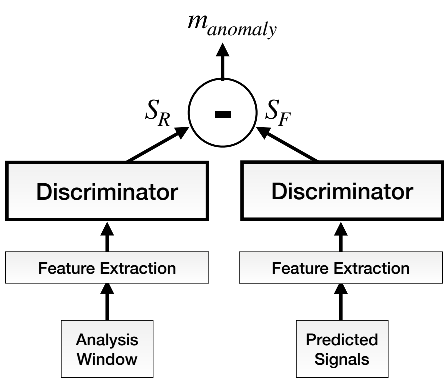
This study proposes the use of conditional GAN for driving anomaly detection using physiological and CAN-Bus data. We use HR, BR, and EDA as the physiological signals, and speed, yaw, pedal angle, brake pressure, steer angle, and steer speed as the CAN-Bus signals. Our formulation is an unsupervised framework that aims to quantify the deviations from the expected driving behaviors. The key idea of the formulation is to predict the vehicle and physiological signals using previous data. These predictions are compared to the actual signals observed during the analysis window. The proposed driving anomaly metric quantifies deviations from the predicted signals, creating a powerful measure to identify unexpected events, which according to our definition, correspond to driving anomalies. A key step in our formulation is to generate the expected vehicle and physiological data. A state-of-the-art generative framework is GAN [4], which has been successfully used in several domains (Sec. 2.3). A conventional GAN uses a noise input to generate the output. For our formulation, we need to constrain the generative model by the data previously observed. This can be achieved with conditional GAN.

Figure 3(a) shows the proposed conditional GAN. The input of the generator (G) is the signals from the previous analysis window and random noise. Its output is the prediction of the signals for the next analysis window. The discriminator (D) takes either the generated or observed signals. Instead of directly using the raw data as input to the generator, we extract statistic features from the data to capture the key aspects of the signals during the analysis window. These statistics are used as features for D. We extract four time domain features for each of the CAN bus data (i.e., maximum, minimum, mean, and standard deviation). For the physiological data, we also extract these temporal statistics, in addition to frequency domain features, corresponding to the energy in the frequency domain covering the following bands: [0-0.04 Hz], [0.04-0.15 Hz], [0.15-0.5 Hz], [0.5-4 Hz], and [4-20 Hz]. The output of D is a discriminative score between (0,1), estimating the probability (S) that the input comes from real signals ($S = 1$) or fake samples generated by G ($S = 0$).

Figure 3(b) illustrates how we estimate the driving anomaly score. First, the generator uses the previous analysis window as a constraint, creating the predicted signal (generator is not shown in Fig. 3(b)). Then, we extract the temporal and frequency statistics from the predicted and actual signals. We input these features into



(a) Training conditional GAN



(b) Estimating anomaly score using discriminator

Figure 3: Unsupervised approach using conditional GAN to identify driving anomaly events. The predicted signals are contrasted with the actual signals. Our approach quantifies these differences using its discriminator.

D to obtain the discriminative score S_F for the predicted signals (fake), and the discriminative score S_R for the actual signals (real). We define our anomaly score $m_{anomaly}$ by estimating the difference between S_F and S_R.

We implement this framework as follows. The generator is designed with a five-layer fully connected *deep neural network* (DNN). The size of the input is 360, where 180 correspond to the first 6-second data, and the remainder 180 correspond to random noise. The five fully connected layers have 180, 60, 18, 60, and 180 nodes, respectively. We implement the discriminator with a four-layer fully connected DNN. The input layer of the discriminator has 51 nodes, where 27 are for the three physiological data, and 24 are for the six CAN-bus data. The four hidden layers are implemented with 51, 34, 17, and 6 nodes, respectively. The output layer has 1 nodes, activated by a Sigmoid function. During training, D and G are trained iteratively for 20 epochs, where the parameters are updated using the ADAM optimizer. All the layers before the output layer use the *exponential linear unit* (ELU) function. The activation function for the output layer of D is the sigmoid function. The analysis window is set to six seconds (i.e., we use six seconds of data to predict the following six seconds of data). Physiological signals do not respond very quickly so it is important to keep the analysis window long enough to capture meaningful patterns.

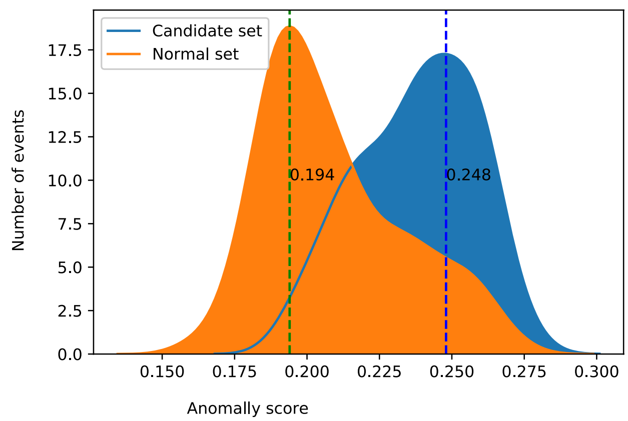


Figure 4: Distribution of anomaly scores $m_{anomaly}$ for segments from the normal and candidate set. The values are higher for the samples in the candidate set.

5 EXPERIMENTAL EVALUATION

This section assesses the benefits of the proposed unsupervised approach with controlled evaluations. The first analysis compares the distribution of the anomaly scores for segments in the candidate and normal sets (Sec. 5.1). The second analysis evaluates the underlying annotations in the DAD overlapping with the segments with high anomaly scores (Sec. 5.2). These results are compared with the annotations of segments selected at random. The third analysis provides perceptual evaluations for segments with high anomaly score and randomly selected segments (Sec. 5.3). Finally, we replicate some of the results using only CAN-Bus data to highlight the benefits of using physiological data 5.4.

5.1 Distribution of Anomaly Scores

The first part of the analysis compares the distributions of the anomaly score for segments in the candidate and normal sets. As explained in Section 3, the segments on the candidate set includes events that are expected to be anomalous (Table 1). The normal set includes segments without any annotation. We consider 400 12-second recordings from each set, using the first six second window to predict the data of the last 6 second windows. Figure 4 shows the distribution of the anomaly scores for both sets. The figure clearly shows that the anomaly scores for segments in the candidate set are often higher than scores for normal segments. This result validates the proposed unsupervised approach, which learns from the data to identify hazard scenarios. Figure 5 shows four examples where our framework assigned high anomaly scores to the videos.

Analyzing cases with unexpected results can provide insights to understand the performance of the system. In Figure 4, we notice that around $m_{anomaly} = 0.25$, which is a relatively high anomaly score, there are some segments from the normal set. In some cases, changes in physiological signals not related to the driving task can trigger our models to predict a high anomaly score. In a segment in

Figure 5: Example of events identified as anomalous. These frames are extracted from videos with high anomaly scores.

Table 2: Assignment of the top 100 events into the normal, candidate, and maneuver sets (Table 1). The table also shows the corresponding assignment for 100 random segments. The assignment is implemented by considering the annotations overlapping with these events.

Events	Normal	Candidate	Maneuver
Top 100	16	9	75
Random 100	59	3	38

the normal set with a high anomaly score, the driver was properly controlling the car, but its BR signal significantly changed from the first six seconds to the last six seconds. This change may be completely unrelated to the driving task (e.g., changing position or suddenly remembering something). While physiological data provide useful information, it is important to remember that not all the physiological changes can be attributed to the driving task.

Some of the recordings in the candidate sets have low anomaly scores. In some cases, on-road bicyclists or pedestrians have been riding or walking on the side of the road for a while, so the drivers had time to prepare, resulting in slight fluctuations on their physiological signals. We can argue that it was the correct decision to assign low anomaly scores to these cases.

5.2 Annotations Overlapping with Segments

The second part of the evaluation explores the annotations provided in the DAD that overlap with recordings with high anomaly scores. We select the 100 recordings with the highest anomaly scores, denoting this group as *Top 100*. For comparison, we randomly select 100 segments, denoting this group as *Random 100*. Depending on the types of annotations spanning the video, these segments are assigned to the normal (i.e., no annotation), candidate (i.e., risky events), or maneuver (i.e., include a driving maneuver) sets.

Table 2 shows the results. Most of the recordings in the top 100 group are included in either the candidate or maneuver set.

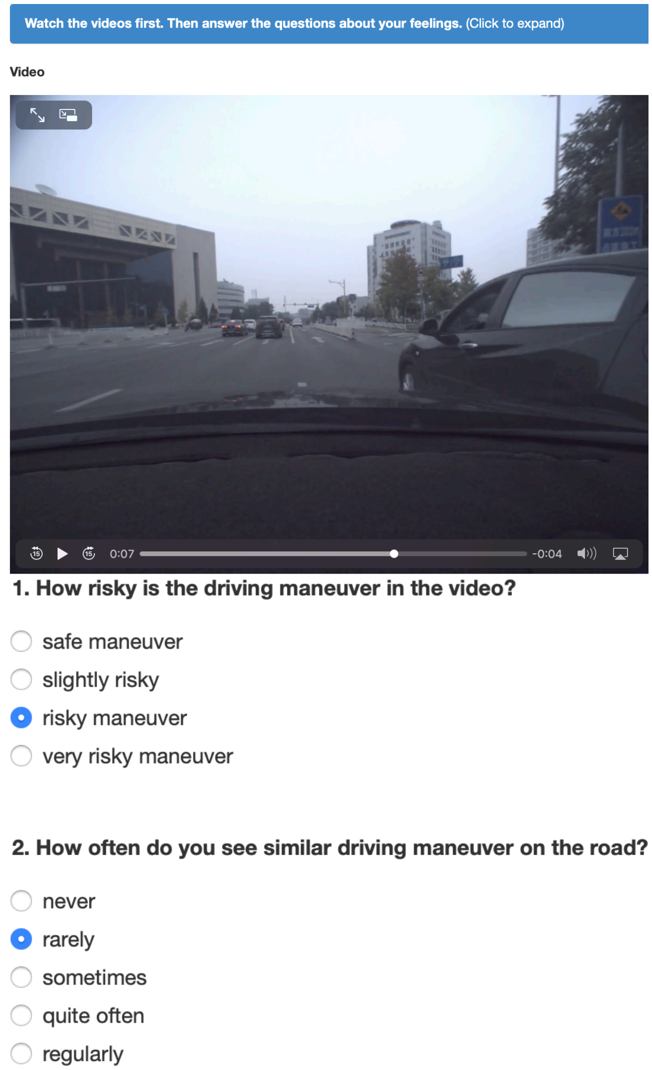
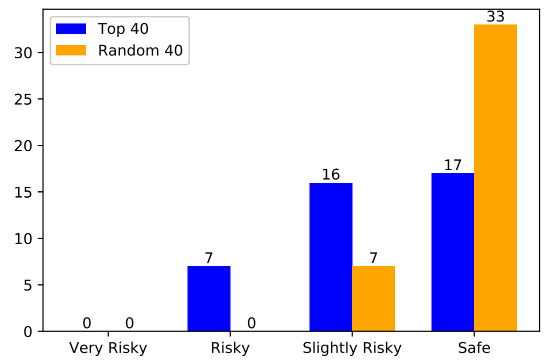


Figure 6: GUI for the perceptual evaluation. The raters are asked to answer the following two questions about the driving scenario, after watching the 12-secs. video.

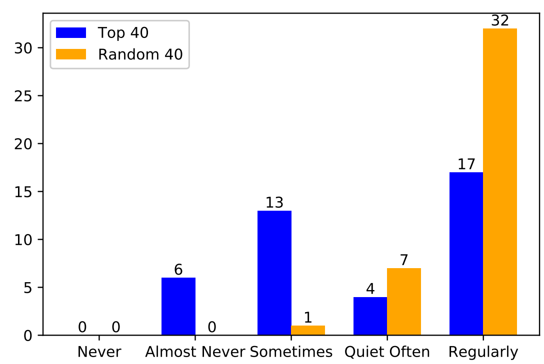
Only 16% of the videos are in the normal set. In contrast, 59% of the recordings on the random 100 group are in the normal group. These results indicate that our supervised method is effective in selecting cases of interest.

5.3 Perceptual Evaluation of Selected Segments

The last analysis evaluates the selected videos with perceptual evaluations. Since perceptual evaluations are expensive and time-consuming, we only considered the top 40 segments with the highest anomaly scores. Each recording is 12 seconds long to provide enough context, including the first six seconds used to predict the signals, and last six seconds. As a baseline, we also evaluated 40 12-second videos that were randomly selected. We asked four raters to watch the videos, where the camera was placed recording the roads. Each rater evaluated 20 videos. We asked questions to assess degree of risk and familiarity of the driving conditions shown in the videos: *how risky is the driving maneuver in the video? (safe maneuver; slightly risky maneuver; risky maneuver; very risky maneuver)* and *how often do you see similar driving maneuver on the roads? (never; almost never; sometimes; quite often; regularly)*. Figure 6 shows the *graphical user interface* (GUI) used for the evaluation.



(a) How risky is the driving maneuver in the video?



(b) How often do you see similar driving maneuver on the roads?

Figure 7: Results of the perceptual evaluation to assess the degree of risk and familiarity of the videos. The figure shows the results for the top 40 segments with the highest anomaly scores, and 40 segments randomly selected.

Figure 7 shows the results of the perceptual evaluation for the two questions. The driving conditions in segments with higher anomaly scores are remarkably more risky and happen less frequently on the roads. For the assessment of risk level, Figure 7(a) shows that 17.5% of the videos identified by our method were labeled as risky. The evaluators perceived some degree of risk in 57.5% of the videos. In contrast, 82.5% of the videos selected at random were perceived as safe. The results in Figure 7(a) show similar patterns for the assessment of familiarity. The evaluators indicated that 15% of the videos selected by our models almost never happen on the roads. Only 42.5% of the selected videos were labeled as driving events that are regularly observed on the roads. In contrast, 80% of the videos selected at random were labeled as events that are regularly observed on the roads. The results of the perceptual evaluation clearly indicate that the proposed unsupervised approach is able to identify anomaly events from the data without defining any rule either manually or through supervised learning.

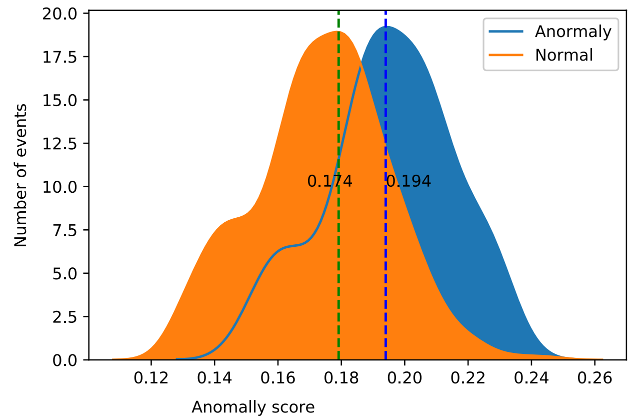


Figure 8: Distribution of anomaly scores $m_{anomaly}$ for segments from the normal and candidate sets. The model only considers features extracted from CAN-Bus data. The overlap between the normal and candidate sets increases, highlighting the benefits of using physiological data.

5.4 Role of Physiological Data

The proposed model relies on physiological and CAN-Bus data. We have shown that physiological data can discriminate driving maneuvers [19]. To evaluate whether physiological data are also useful for anomaly driving detection, we reimplemented the network using only CAN-Bus features. In particular, we consider the results with the distribution of anomaly scores in Figure 4 (Sec. 5.1) and the assignment of the top 100 events into the normal, candidate, and maneuver sets in Table 2 (Sec. 5.2).

Figure 8 shows that using only CAN-Bus features increases the overlap between the distributions of the normal and candidate sets. With CAN-Bus and physiological features, the modes of the distributions are 0.194 (normal) and 0.248 (candidate). The separation between the modes is 0.054. With only CAN-Bus features, the modes of the distributions are 0.174 (normal) and 0.194 (candidate). The separation between modes is only 0.02.

Table 3 shows that the top 100 set includes more cases labeled as "normal" when we only use CAN-Bus features (Normal: 16 $\rightarrow$ 31; Candidate: 9 $\rightarrow$ 6; Maneuver: 75 $\rightarrow$ 53). These results illustrate that the discrimination power of the model is improved by adding physiological features. As a particular case, we describe the scenario displayed on Fig 6. The driver slowed down the vehicle to turn right, but a black car on the right lane rushed forward from the right rear. This sudden change did not cause a big change on CAN-Bus

Table 3: Assignment of the top 100 events into the normal, candidate, and maneuver sets. The table also shows the corresponding assignment for 100 random segments. The model was implemented with only CAN-Bus features.

Events	Normal	Candidate	Maneuver
Top 100	31	6	53
Random 100	57	4	39

data, but caused a noticeable drop on the driver's breath rate. The anomaly score was 0.231 with physiological features, and 0.208 without physiological features.

6 CONCLUSIONS

This study proposed a multimodal unsupervised driving anomaly detection system using a conditional GAN. The proposed approach creates predictions of physiological signals and CAN-Bus data, conditioning the models by the previous analysis window. The predictions are contrasted with the actual data quantifying the differences. This approach is implemented by subtracting the scores provided by the discriminator when the input features are either the actual or predicted data. This approach is effective in detecting events that deviates from the predicted driving behaviors, creating solutions that do not depend on predefined rules set with either ad-hoc thresholds or supervised methods.

The proposed model was trained and evaluated on a novel driving dataset (DAD) using 48 hours of naturalistic recordings. By considering the annotations provided in the dataset, we observed that the proposed anomaly scores are higher for recordings annotated with risky events than for the recordings without any annotation. Our perceptual evaluation demonstrated that the proposed unsupervised method is able to identify recordings that are perceived as riskier and more uncommon than randomly selected driving recordings. By considering the physiological and CAN-Bus data, the system can detect driving anomalies signaled by changes in the driver's mental state or unexpected driving maneuvers.

A limitation of this study is that our detection algorithm can identify anomalies only when the driver reacts to them. Our approach cannot detect anomalies that were missed by the driver (i.e., when a driver did not react to events or objects affecting the driving conditions, neither physically nor mentally). Another important area of improvement is the normalization of physiological data. Physiological responses vary from drivers to drivers, depending on their expertise, habits and driving conditions. In this study, we normalized the physiology data per session to compensate for the differences between drivers and road situations. However, we will need to develop an algorithm that normalize the driver's physiological data to apply the proposed approach in practical applications.

This study opens interesting research opportunities to improve the proposed approach. We only considered physiological and CAN-Bus data. In addition, we can also incorporate other information such as the results from vision-based object detection systems applied to the road. Another potential modification of the system is to consider conditional GANs implemented with *recurrent neural networks* (RNNs) to capture temporal information. The current approach is implemented with fully connected layers, where the features of the discriminator are statistics from physiological and CAN-Bus data. Instead, we can also implement the system with *convolutional neural networks* (CNNs) to obtain better feature representations learned directly from the raw signal.

7 ACKNOWLEDGMENTS

This work was supported by Honda Research Institute USA.

REFERENCES

[1] Zhongyang Chen, Jiadi Yu, Yanmin Zhu, Yingying Chen, and Minglu Li. 2015. D3: Abnormal driving behaviors detection and identification using smartphone sensors. In *2015 12th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. 524–532. https://doi.org/10.1109/SAHCN.2015.7338354

[2] Bo Dai, Sanja Fidler, Raquel Urtasun, and Dahua Lin. 2017. Towards Diverse and Natural Image Descriptions via a Conditional GAN. In *The IEEE International Conference on Computer Vision (ICCV)*.

[3] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. 2013. Vision meets Robotics: The KITTI Dataset. *International Journal of Robotics Research (IJRR)* (2013).

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 2672–2680. http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

[5] John H.L. Hansen, Carlos Busso, Yang Zheng, and Amardeep Sathyanarayana. 2017. Driver Modeling for Detection and Assessment of Driver Distraction: Examples from the UTDrive Test Bed. *IEEE Signal Processing Magazine* 34, 4 (July 2017), 130–142. https://doi.org/10.1109/MSP.2017.2699039

[6] Minamitani Haruyuki. 1997. Measurement of Physiological Signals of Stress State. *The Journal of the Institute of Electronics,Information and Communication Engineers* 80, 7 (Jul 1997), 754–757.

[7] Jennifer A. Healey and Rosalind W. Picard. 2005. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on intelligent transportation systems* 6, 2 (June 2005), 156–166. https://doi.org/10.1109/TITS.2005.848368

[8] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. 2017. Image-To-Image Translation With Conditional Adversarial Networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[9] Dan Li, Dacheng Chen, Jonathan Goh, and See-Kiong Ng. 2018. Anomaly Detection with Generative Adversarial Networks for Multivariate Time Series. *CoRR* abs/1809.04758 (2018).

[10] Nanxiang Li and Carlos Busso. 2015. Predicting Perceived Visual and Cognitive Distractions of Drivers with Multimodal Features. *IEEE Transactions on Intelligent Transportation Systems* 16, 1 (February 2015), 51–65. https://doi.org/10.1109/TITS.2014.2324414

[11] Nanxiang Li and Carlos Busso. 2016. Detecting Drivers' Mirror-Checking Actions and its Application to Maneuver and Secondary Task Recognition. *IEEE Transactions on Intelligent Transportation Systems* 17, 4 (April 2016), 980–992. https://doi.org/10.1109/TITS.2015.2493451

[12] Nanxiang Li and Carlos Busso. 2017. Driver Mirror-Checking Action Detection. In *DSP for In-Vehicle Systems and Safety*, H. Abut, J.H.L. Hansen, G. Schmidt, K. Takeda, and H. Ko (Eds.). DeGruyter, 55–76. https://doi.org/10.1515/9781501504129-005

[13] Nanxiang Li, Jinesh J. Jain, and Carlos Busso. 2013. Modeling of Driver Behavior in Real World Scenarios using Multiple Noninvasive Sensors. *IEEE Transactions on Multimedia* 15, 5 (August 2013), 1213–1225. https://doi.org/10.1109/TMM.2013.2241416

[14] Nanxiang Li, Teruhisa Misu, and Alexandre Miranda. 2016. Driver behavior event detection for manual annotation by clustering of the driver physiological signals. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. 2583–2588. https://doi.org/10.1109/ITSC.2016.7795971

[15] Yi lu Murphey, Dev S. Kochhar, Paul Watta, Xipeng Wang, and Tianyu Wang. 2015. Driver Lane Change Prediction Using Physiological Measures. *SAE Int. J. Trans. Safety* 3 (April 2015), 118–125. https://doi.org/10.4271/2015-01-1403

[16] Lucas Malta, Chiyomi Miyajima, and Kazuya Takeda. 2009. A Study of Driver Behavior Under Potential Threats in Vehicle Traffic. *IEEE Transactions on Intelligent Transportation Systems* 10, 2 (June 2009), 201–210. https://doi.org/10.1109/TITS.2009.2018321

[17] Teruhisa Misu and Yiting Chen. 2018. Toward Reasoning of Driving Behavior. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. 204–209. https://doi.org/10.1109/ITSC.2018.8569250

[18] Morimichi Nishigaki, Ryota Mose, Osamu Takahata, Hideki Imafuku, and Hironori Aoygai. 2018. Quantitative evaluation on mental worklosd reduction for hands free driving. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. 230–235. https://doi.org/10.1109/ITSC.2018.8569638

[19] Y. Qiu, T. Misu, and C. Busso. 2019. Analysis of the Relationship Between Physiological Signals and Vehicle Maneuvers During a Naturalistic Driving Study. In *Intelligent Transportation Systems Conference (ITSC 2019)*. Auckland, New Zealand.

[20] Vasili Ramanishka, Yi-Ting Chen, Teruhisa Misu, and Kate Saenko. 2018. Toward Driving Scene Understanding: A Dataset for Learning Driver Behavior and Causal Reasoning. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018)*. Salt Lake City, UT, USA, 7699–7707. https://doi.org/10.1109/CVPR.2018.00803

[21] Najmeh Sadoughi and Carlos Busso. 2019. Speech-Driven Expressive Talking Lips with Conditional Sequential Generative Adversarial Networks. *IEEE Transactions on Affective Computing* To appear (2019). https://doi.org/10.1109/TAFFC.2019.2916031

[22] Donald Selmanaj, Matteo Corno, and Sergio M. Savaresi. 2014. Accelerometer-based data-driven hazard detection and classification for motorcycles. In *2014 European Control Conference (ECC)*. 1687–1692. https://doi.org/10.1109/ECC.2014.6862549

[23] Adela C. Timmons, Theodora Chaspari, Sohyun C. Han, Laura Perrone, Shrikanth S. Narayanan, and Gayla Margolin. 2017. Using Multimodal Wearable Technology to Detect Conflict among Couples. *Computer* 50, 03 (mar 2017), 50–59. https://doi.org/10.1109/MC.2017.83

[24] Hongyang Zhao, Huan Zhou, Canfeng Chen, and Jiming Chen. 2013. Join driving: A smart phone-based driving behavior evaluation system. In *2013 IEEE Global Communications Conference (GLOBECOM)*. 48–53. https://doi.org/10.1109/GLOCOM.2013.6831046

[25] Yang Zheng and John. H. L. Hansen. 2016. Unsupervised driving performance assessment using free-positioned smartphones in vehicles. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. 1598–1603. https://doi.org/10.1109/ITSC.2016.7795771