Is It Safe to Drive? An Overview of Factors, Metrics, and Datasets for Driveability Assessment in Autonomous Driving

Junyao Guo, Unmesh Kurup, and Mohak Shah

***Abstract*— With recent advances in learning algorithms and hardware development, autonomous cars have shown promise when operating in structured environments under good driving conditions. However, for complex, cluttered, and unseen environments with high uncertainty, autonomous driving systems still frequently demonstrate erroneous or unexpected behaviors that could lead to catastrophic outcomes. Autonomous vehicles should ideally adapt to driving conditions; while this can be achieved through multiple routes, it would be beneficial as a first step to be able to characterize *driveability* in some quantified form. To this end, this paper aims to create a framework for investigating different factors that can impact driveability. Also, one of the main mechanisms to adapt autonomous driving systems to any driving condition is to be able to learn and generalize from representative scenarios. The machine learning algorithms that currently do so learn predominantly in a supervised manner and consequently need sufficient data for robust and efficient learning. Therefore, we also perform a comparative overview of 54 public driving datasets that enable learning and publish this dataset index at https://sites.google.com/view/driveability-survey-datasets. Specifically, we categorize the datasets according to the use cases and highlight the datasets that capture the complicated and hazardous driving conditions, which can be better used for training robust driving models. Furthermore, by discussions of what driving scenarios are not covered by the existing public datasets and what driveability factors need more investigation and data acquisition, this paper aims to encourage both targeted dataset collection and the proposal of novel driveability metrics that enhance the robustness of autonomous cars in adverse environments.**

***Index Terms*— Autonomous driving, driveability metric, machine learning, public dataset, risk assessment, traffic hazards.**

I. Introduction

DESPITE testing autonomous cars in highly controlled settings, these cars still occasionally fail in making correct decisions, often with catastrophic results.[1] There have been several accidents reported recently [1] due to failure of the autonomous capability of these cars. According to the accident records, the failures are most likely to happen in complex or unseen driving environments. The fact remains that while autonomous cars can operate well in controlled or structured environments such as highways, they are still far from reliable when operating in cluttered, unstructured or unseen environments [2].

To adapt autonomous driving systems to all types of driving conditions, it would be beneficial to first characterize the *Driveability* of a scene.[2] This can then lead to addressing two core issues of autonomous driving or advanced driver-assistance systems (ADAS): 1) policy learning for driver control hand-off; and 2) incorporation of *driveability* in the autonomous vehicle's decision making, planning and testing. These two different application fields also suggest that *driveability* could be quantified in different forms, either as a single metric or a composition of metrics. For example, with ADAS and current Level 2 or 3 autonomy, a scene can be simply defined as *driveable* if the car can operate safely in autonomous mode. When a non-driveable scene is detected, the autonomous car can hand over control to the human driver in a timely manner [4]. However, in the long term where Level 4 or 5 autonomy is targeted, it is not possible to hand over control to the driver. This restriction means that a richer representation of driveability is needed; one that is informative enough for the car to take proactive measures to prevent catastrophic failure. For instance, an autonomous car can selectively use sensors in scenes with high driveability whereas request further support from cloud computing for detailed analysis of less-driveable scenes [5]. In scenes with particularly low driveability, contingency plans might need to be executed such as deceleration or even making a full stop. The driveability information could also be useful when processed offline. For example, it can be used to build a safety map that not only guides road users to plan alternate routes that are safer, but also identifies unsafe areas that need repair [6]. Such information can also be used by insurance companies to quantify operating risk [7], [8].

The concept of driveability is not new. It has been proposed for measuring road conditions [9] and modeling human driver performance [10]. However, there exists no unified concept of driveability in the domain of autonomous driving. Related usages of driveability include "driveability map" which is a map divided into cells categorized as driveable or non-driveable for motion planning [11]–[14], and "object

Manuscript received December 6, 2018; revised April 23, 2019; accepted June 21, 2019. Date of publication July 12, 2019; date of current version July 29, 2020. The Associate Editor for this paper was H. G. Jung. *(Corresponding author: Junyao Guo.)*

The authors are with the Advanced AI Group, LG Silicon Valley Lab, Santa Clara, CA 95054 USA (e-mail: junyao.guo@lge.com; unmesh.kurup@lge.com; mohak.shah@lge.com).

Digital Object Identifier 10.1109/TITS.2019.2926042

[1]These apply to autonomous vehicles in general.

[2]We adopt the definition of a *scene* proposed in [3]; i.e., "A scene describes a snapshot of the environment including the scenery and dynamic elements, as well as all actors' and observers' self-representations, and the relationships among those entities."

1524-9050 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

driveability" which determines if an object can be driven over without causing damage to the vehicle [15]. The concept that is most similar to that considered in this paper is "scene driveability" proposed in two recent studies [4], [16]. Scene driveability measures how easy a scene is for an autonomous car to make full decisions on steering angles and accelerations or to perform a specific task such as lane changing. However, none of these studies reason about what makes a scene driveable or non-driveable at a high level in the context of autonomous driving.

While it may not be difficult for human drivers to tell whether an environment is safe, it is far from obvious for an autonomous car because the driveability is affected by a variety of factors. Therefore, we first provide one taxonomy of those factors including both environmental conditions and road user behaviors, and particularly review representative approaches that handle the risks originating from each factor. Then, established performance metrics for driveability assessment are summarized. The development of both robust methods and novel metrics depends, in part, on having access to large-scale naturalistic and diverse driving datasets. This will have consequence not just in understanding of driveability conditions and developing novel quantification metrics, but will also assist subsequent validation and verification as well as incorporation of driveability aspects in autonomous driving policy. Therefore, as a major contribution of this paper, a comprehensive and comparative study of publicly available datasets for autonomous driving research is presented. To provide practical guidance on when to use which dataset, we categorize the datasets according to autonomous driving tasks. More importantly, we highlight the datasets that capture low-driveability scenes using which models could be trained to better handle traffic hazards and risks.

By reviewing existing literature and datasets from a driveability perspective, this paper accomplishes the following:

- This paper proposes the concept of "Driveability" in the context of autonomous driving, and identifies factors that affect driveability based on academic research, government reports and industrial practices. There is no systematic study on how well current autonomous vehicles can handle these factors, what metrics are used to measure driveability, and what dataset can be used to study each factor. This paper fills this gap by organizing related literatures from the driveability perspective, and aims to provide researchers with practical references on improving the robustness of autonomous driving models.
- This paper presents a learning pipeline for driveability assessment and establishes the link between factors, metrics, and datasets. By an in-depth review of related literatures along this learning pipeline, this paper identifies multiple challenges in handling various driveability factors, designing principled driveability metrics, and developing autonomous driving algorithms where realistic data is not sufficient or even available. Discussions made in this paper will encourage research to address these challenges and thereby enhance the reliability of autonomous driving systems.
- This paper provides an up-to-date overview of publicly open driving datasets, with focus on how to use these datasets. Different from existing dataset surveys, this paper provides a more practical comparison of datasets and classifies datasets according to both autonomous driving tasks and driveability factors. This provides not only suggestions on dataset selection for a specific task, but also insights into which driving scenario needs more targeted dataset collection and annotation. To the best of our knowledge, this is the most comprehensive dataset overview that provides practical guidance on when to use what dataset for autonomous driving studies.

The rest of the paper is organized as follows: Section II presents the factors contributing to driveability and their related studies and challenges. Section III introduces existing metrics used for driveability assessment and discusses their limitations. Section IV carries out the study of existing public driving datasets. Section V and Section VI propose approaches that enable learning when data is scarce. Finally, Section VII concludes the paper.

II. Driveability Factors

The driveability of a scene is greatly affected by environmental conditions such as weather, traffic flow, road condition, obstacles, and so on, which are explicit factors that can be directly perceived from the environment. However, environmental factors alone are not sufficient for driveability assessment. At times, potential risks can only be identified if the intent and interactions of road users are understood [17], [18], which is implicit information that needs to be inferred from observation. Therefore, in this section, we present both explicit and implicit factors that contribute to driveability. These factors are identified based on various resources including driveability studies in other transportation system research fields [9], [10], U.S. Department of Transportation reports [19], [20], and industrial standards for road risk estimation [21].

Table I provides an overview of the driveability factors, summarizing the hazardous scenarios and difficulties associated with each factor, and classifying related literatures and datasets according to the factors. The hazardous scenarios denote the conditions that generally lower the driveability, which are more difficult to handle compared to good or controlled driving conditions assumed in most autonomous driving studies. The dataset column in Table I includes the most representative datasets that cover hazardous scenarios of each factor, which will be discussed in details in Section IV.

Handling each factor presents its own research challenge and existing approaches are still far from delivering robust solutions that incorporate all factors. To maintain focus over a large research landscape, we will point to surveys on generic studies of these factors and only detail representative works that focus on scenarios with high complexity, dynamics and uncertainty.

A. Explicit Factors

1) Weather: Extreme weather such as *fog*, *heavy rain* and *snow* can significantly impair road visibility and therefore

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

TABLE I
SUMMARY OF DRIVEABILITY FACTORS, RELATED REFERENCES AND DATASETS

Factor	Hazadous scenarios	Difficulties in handling	References	Datasets
Explicit/Environmental factors				
Weather	fog, rain, snow	impair visibility and cause difficulty in vision-based tasks; rain can cause reflection on wet roads	[22]–[24]	AMUSE, CMU, HCI Challenging Stereo, Oxford RobotCar, UTBM
Illumination	shadow, night, dawn/dusk, directly facing the sun	large variations in light intensity; nighttime pedestrian detection	[25]–[31]	ApolloScape, NightOwls
Road geometry	intersections, roundabouts, rural or unmapped areas	different traffic patterns, markings, signs and road user behaviors at different road geometries	[32]–[34]	Berkeley DeepDrive, Málaga, UTBM Robocar
Road condition	road damage, uneven and rough surfaces, road construction	affect road appearances; lack of labeled data	[35]–[37]	CCSAD, IDD, Road Damage, Highway Workzones
Lane marking	broken/missing markings, irregular lane/road shapes	non-uniformity and frequent changes in lane appearances; visual cue is insufficient	[38]–[41]	ApolloScape, Argoverse, Berkeley DeepDrive, CULane, KAIST Multi-Spectral, TuSimple
Traffic condition	high speed limit, heavy flows, potential accidents	lack of realistic dataset	[42]	ApolloScape, Berkeley DeepDrive
Objects	small size, heavy occlusion, large truncation, sudden/unexpected appearance	high heterogeneity in appearances; hard to identify unseen objects	[43]–[50]	Bosch Small Traffic Lights, LostAndFound
Implicit/Behavioral factors				
Vehicle behaviors	overtaking, lane change, rear-ending, speed-driving, failure to obey traffic laws	real-time prediction at high speed driving; large volume of vehicles in cluttered urban areas	[44], [51]–[55]	Argoverse, Brain4Cars, UAH
Pedestrian behaviors	crossing, inattentiveness, failure to obey traffic laws	abrupt changes; affected by numerous factors such as age, gender, social norms, culture, vehicle distance, etc	[56]	JAAD
Driver behaviors	distraction, drowsiness	emergency situations	[57]–[60]	Dr(eye)ve, JAAD
Motorcyclist/ Bicyclist behaviors	crossing, traveling in wrong direction	lack of datasets	[61]	Daimler Pedestrian, NightOwls

cause difficulty on many vision-based tasks such as road detection and object tracking. For instance, deep neural network (DNN) models are known to behave erroneously under adverse weather conditions [62]. Existing works that handle adverse weathers mainly focus on road detection. Two major approaches are pursued, namely extraction of more weather-invariant vision features [22], [23] and sensor fusion using both camera and LiDAR data [24]. However, it is not well understood how various weather conditions affect the robustness of models that address other perception tasks other than road detection.

2) Illumination: Variations in illumination caused by *time of day (dusk/dawn/night)*, *landscape (shades)* and *light sources* pose different challenges for environment perception. References [25]–[27] propose various illumination invariant transforms to improve the robustness of visual perception of autonomous cars under different directions and intensities of the light sources. Night scenes require special handling due to the large variations in light intensity caused by inhomogeneous light sources and their reflections [63]. For nighttime vehicle detection, a common approach is to use vehicle lights as detectable features [28], [29]. However, nighttime pedestrian detection is much harder. It is shown in [63] that the accuracies of six state-of-the-art vision-based pedestrian detectors drop drastically when trained and evaluated on night scenes. Due to the limitation of visual features under poor lighting, a couple of recent studies [30], [31] propose methods and datasets that incorporate additional thermal or far infrared data for more accurate nighttime pedestrian detection.

3) Road Geometry: Intersections and *roundabouts* are more difficult to drive through compared to straight highway roads. As pointed out in [64], a significant number of accidents involve intersections. Consequently, intersection is the most studied road geometry, where [32] makes an overview of recent studies on road users behaviors at intersections. Another notable work is [33], which uses different combinations of visual cues such as scene flow, semantic labels, vanishing points, etc for intersection scene understanding including object detection, driveable area estimation, street orientation and lane detection. Compared to intersections, roundabouts are even more challenging given the stringent time constraints on yielding and merging maneuvers. Reference [34] reviews related studies and proposes an action planning method to enable an autonomous car to merge into a roundabout. While there are many studies focusing on intersections and roundabouts, almost all consider urban scenarios. In contrast, for rural and unmapped areas, the road geometry and its impact on autonomous driving are not well understood.

4) Road Condition: Potentially hazardous road conditions include *road damage*, *uneven and rough surfaces*, *road construction*, etc. However, due to the fact that hazardous road conditions are relatively uncommon, a major barrier for identifying these conditions is lack of labeled data. Reference [35] introduces a large-scale road damage dataset and implements

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

a convolutional neural network (CNN)-based method to detect eight categories of road damage. Road construction can affect multiple road appearances including road geometry and traffic signs as well as driving condition. Usually, there are workers present around a construction site that also need particular attention. Therefore, the boundary of road construction needs to be accurately identified and potential hazards there need to be specifically handled. To this end, [37] introduces a highway work zone dataset for identification of construction site boundary and driving condition change on highways. Terrain roughness is also important for vehicle control because rough terrain induces shock especially at high speed. Survey [36] reviews methods that estimate terrain traversability for unmanned ground vehicles which could potentially be extended to autonomous cars. However, there is no public dataset that captures and labels terrain conditions.

5) Lane Marking: Lane markings provide important visual cues for lane or road detection, which is an active research area where a classic approach is to fit a geometric model to the visual features extracted from lane markings [38]. However, there exist many unstructured roads in both urban and rural areas with *broken or missing lane markings* or *irregular lane shapes*, which makes road detection difficult because it cannot solely rely on lane marking recognition. References [39]–[41] are some representative studies that focus on improving the robustness of road detection methods for roads with no markings or irregular lane shapes. A common idea in these studies is to synergize multiple visual cues such as both monocular and stereo vision or fuse multiple sensor inputs from camera, infrared and LiDAR.

6) Traffic Condition: In general, highway and urban scenarios require specified learning models due to their different traffic characteristics. Particular attention needs to be paid to situations with extremely *high speed limit*, *heavy flows* and *potential accidents*. However, majorly due to lack of datasets, it is not well understood whether/how the autonomous car's behavior is affected by traffic conditions or when an accident is likely to happen. To this end, [42] presents a near-miss accident dataset and uses a quasi-recurrent neural network (RNN) to predict accidents.

7) Static and Dynamic Objects: Object detection and tracking is the most studied topic with surveys [43]–[47] reviewing methods and benchmarks on detection of static obstacles, vehicles, pedestrians and cyclists. However, existing methods still demonstrate high error when encountering unseen or hard-to-identify objects especially those of *small size*, *heavy occlusion* and *large truncation* [50]. Towards detection of uncommon and unexpected obstacles, [48] constructs a dataset with small obstacles such as lost cargos and proposes a Bayesian fusion framework that combines CNN with a stereo-system for semantic label prediction. Another work [49] focuses on obstacles that have thin structures such as cables and tree branches, and proposes an edge-based visual odometry technique for detection. While these studies cover several types of uncommon obstacles, there are many more that need investigation such as animals, flying hindrances, sudden obstruction, etc. Furthermore, many of these obstacles (like fallen trees) can cause sudden and unexpected changes to the a-priori map required for vehicle localization. These changes can potentially impair localization accuracy and consequently hinder effective planning of autonomous cars.

B. Implicit Factors

Implicit factors consist of behaviors and intent of road users interacting with the autonomous car. An interesting survey [57] exhaustively reviewed studies on understanding, modeling, and predicting human agents in three domains, namely, inside the vehicle cabin, around the vehicle, and inside surrounding vehicles. Reference [65] also provides a literature review on interaction between autonomous cars and other road users. We will not repeat the findings presented in these surveys, but only highlight risky behaviors of vehicles, pedestrians, drivers, motorcyclists and bicyclists, respectively.

1) Vehicle Behaviors: Potentially hazardous vehicle behaviors include *overtaking*, *lane change*, *rear-ending*, *speed-driving* and *failure to obey traffic laws*, where the first three are studied the most with related works summarized in [44], [51], [52]. Vehicle behavior prediction can be hard at high-speed or cluttered urban scenarios. Under these considerations, [53] reviews studies on trajectory planning and tracking for high-speed autonomous overtaking, whereas [54], [55] propose scenario-adaptive and intention-aware vehicle trajectory estimation approaches for challenging urban scenarios such as unsignalized urban intersections. Note that vehicle behaviors are tightly connected to the driver's condition. Hazardous vehicle behaviors can also occur when their respective drivers are inexperienced, intoxicated, or limited in physical ability to pay careful attention to everyone on the road.

2) Pedestrian Behaviors: Pedestrians are the most vulnerable road users. Most accidents happen when a pedestrian is *crossing*, and many seem to result from pedestrians' *inattentiveness* or *failure of compliance with law*. Reference [56] gives a comprehensive overview of factors, methods and challenges on pedestrian behavior studies with both the absence and presence of autonomous vehicles. As summarized in [56], there are two common approaches to predict pedestrian intention. One is to treat it as a dynamic object tracking problem and calculate future pedestrian's trajectory, whereas the other is to cast it as a classification problem, e.g., classifying a pedestrian's behavior into "crossing" or "not crossing". However, most methods are only based on the pedestrian's poses and positions, which are subject to error because a pedestrian's intent or behavior can change abruptly. It is argued that other factors should be taken into account for pedestrians' behavior estimation including demographics such as age and gender, dynamic factors such as vehicle speed and distance, as well as social norms and culture. It remains an open question how these factors are interrelated and influential in understanding pedestrians' intent.

3) Driver Behaviors: For partial or high automated vehicles, driver's availability is still required at times when the autonomous car is incapable of making reliable decisions and requests to handover control. Driver behavior analysis is the most mature field with a myriad of studies on driver's activity, intent, alertness, skill and style, and so on [57]. Among these aspects, driver *distraction* and *drowsiness* are two main

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

reasons for traffic accidents. Classic approaches for driver distraction-and drowsiness-detection include using visual features such as facial expression and eye movement, non-visual features such as physiological signals and car dynamics, or hybrid features that combine both. Surveys [58]–[60] review these approaches where they show that hybrid measures generate fewer false alarms and higher recognition rate. More investigation is needed on handling emergency situations such as sudden driver impairment under medical conditions.

4) Motorcyclist/Bicyclist Behaviors: Compared to other groups of road users, the models and methods developed for bicyclists/motorcyclist behavior analysis are far more limited. This is largely due to the lack of datasets that capture and label the trajectories and behaviors of motorcyclists and bicyclists. Report [61] shows that the majority of accidents happen when a bicycle appears in front of a vehicle and the driver failed to brake in time due to either obstructed view or the bicyclist traveling in the wrong direction. Substantial efforts are needed for trajectory estimation and intent prediction for motorcyclists and bicyclists.

C. Interrelationship of Driveability Factors

In most cases, the driveability factors are interrelated where the change in one factor could affect the others significantly. For example, how road users behave greatly depend on explicit factors such as weather, road condition, road geometry, and traffic condition. While there are many studies on behavior analysis under different driving contexts, those contexts are generally limited to road geometry such as intersection and traffic condition such as highway traffic versus urban traffic [44], [52]. The interrelationship among other factors are under-exploited.

A more challenging issue arises from the interaction of different groups of road users. The biggest limitation of existing research on behavior analysis is that most studies only focus on one type of road user (vehicles or pedestrians only). Only a few works investigate the *joint attention* of multiple road users including autonomous cars appearing in the same scene [66], [67]. Joint attention in autonomous driving is a complicated problem which involves issues from biological, social and algorithmic perspectives and requires methods for multiple tasks including object detection and tracking, pose estimation and intent prediction [67].

As pointed out in [68], autonomous cars should ideally behave in a way that is comprehensible to humans. They should communicate effectively with pedestrians and cyclists, and react safely to any unpredicted human behaviors. Beyond interaction with pedestrians, it is also important to plan for autonomous cars that account for effects on human drivers because they are expected to share the road for the coming decades. In fact, the cause of many accidents involving autonomous cars could be attributed to human drivers expecting autonomous cars to behave differently [69]. In this direction, one pioneering work [70] uses an inverse reinforcement learning approach to optimize the planning for autonomous vehicles that takes into account the effects on human drivers, which makes the autonomous vehicle more efficient and communicative. However, this study was only carried out on simple simulated scenarios, and more investigation is needed about its applicability in real-world situations.

Fig. 1. Overview of the learning pipeline for driveability assessment.

III. Driveability Metrics

Currently, there is no unified performance metric for assessing the driveability of a scene, because driveability assessment involves many tasks in perception and behavior analysis, and the performance metric for each task is tightly related to the underlying model used. For a better understanding of the relationship among the autonomous driving tasks, driveability factors and related metrics, we first illustrate a learning pipeline for driveability assessment in Fig. 1. This learning pipeline is built based on the autonomous driving frameworks released by both research institutes [71], [72] and industrial companies [73], [74].

There are mainly two approaches to achieve driveability assessment, namely, a modular approach and an end-to-end approach. The modular approach feeds the sensor data first into perception modules to extract, recognize and analyze explicit/environmental factors. Environment perception is mainly computer vision-based, which usually includes low-level perception tasks such as stereo/3D vision and optical flow, and high-level perception tasks such as object detection and tracking, lane detection and semantic segmentation. We refer interested readers to [43] for a comprehensive overview of these perception tasks. Then, based on robust environment perception, behavior analysis can be carried out to identify and analyze implicit/behavioral factors, including recognizing vehicle maneuvers and predicting the intent and trajectories of all traffic participants [44]. Finally, the risks can be estimated by incorporating results from both environment perception and behavior analysis. In contrast, an end-to-end approach does not explicitly capture individual driveability factors, but directly assesses the driveability of a scene based on sensor inputs.

The most relevant metrics established in existing research are the metrics for scene driveability [4], collision-based risk [75] and behavior-based risk [75], which are associated

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

with the end-to-end module, the risk estimation module and the behavior analysis module in the learning pipeline, respectively. Note that in the modular approach, there are various low-level metrics that evaluate each driveability factor such as accuracy for object detection, which are not the focus of this survey. In other words, the driveability metric should be a high-level estimator of the overall driving condition derived from these lower-level measures and potentially other sources. While the end-to-end approach does not explicitly analyze each driveability factor, a good driveability metric for this approach should capture the variations in the driveability factors as well.

In this section, we present the design methodology underlying these existing metrics, with the purpose of encouraging the proposal of novel metrics for driveability assessment. While risk assessment metrics are well established and accepted in studies on ADAS, a metric for scene driveability has only been proposed recently and it is aimed at end-to-end driving policy learning. For risk assessment, we follow the discussions made in [75], and we refer interested readers to [75] for details on the works that utilize these metrics.

A. Scene Driveability

In [4], scene driveability is defined by how easy a scene is for an autonomous car to navigate and a scene driveability score is used to measure how likely the car will fail. An end-to-end approach is used to calculate this score. Specifically, an end-to-end driving policy learning model using CNN and long-short-term-memory (LSTM) is first trained to predict driving maneuvers including velocity and steering angle. Then the scene driveability score is calculated based on the discrepancies between the predictions made by the trained driving model and the ground-truth maneuvers. If the score is lower than some manually chosen threshold, the scene is considered "Hazardous". Otherwise, the scene is considered "Safe". Once all the scenes are labeled safe or hazardous, another CNN model is trained to predict whether a new scene is safe or hazardous. A similar approach is used in [16], which uses a bi-directional RNN to classify scenes as safe or unsafe for performing a lane change. However, the limitation of such metrics is that they are defined purely from the model prediction outcome, which is highly model dependent. Furthermore, the end-to-end approach barely provides any insight into what makes a scene hazardous.

B. Collision-Based Risk

There are two basic metrics for collision risk computation, namely, binary risk indicator and probabilistic risk indicator. The former only predicts whether or not a collision will happen in the near future, while the latter represents risk score by a probability calculated based on current states, event, choice of hypothesis, future states and damages [76], [77]. Conventional methods that calculate collision-based risk first predict the potential future trajectories for moving entities and then detect collisions between each pair of trajectories [75]. Newer approaches use deep predictive models to predict whether a collision will occur in the future directly from videos and other sensor data [78].

Another widely used indicator is Time-To-X (TTX), where X refers to a relevant event in the course towards collision. The most standard TTX indicator is Time-To-Collision (TTC), which measures the time remaining before the collision occurs and provides clues on whether the car should send a warning to the driver or directly perform an action. Recent study [79] also proposed a worst-time-to-collision (WTTC) metric with the purpose of selecting most critical objects and situations out of a typical test drive to reduce the amount of data saved. An object or situation is considered less critical if its associated WTTC is large. However, the calculation of TTC in most studies relies on simple assumptions of vehicle status and trajectories, which may be difficult to adapt to real-world driving scenarios.

C. Behavior-Based Risk

Behavior-based risk estimation is usually cast as a binary classification problem, where "nominal behaviors" are learned from data and then "dangerous behaviors" are detected. Nominal behaviors are defined based on acceptable speeds, traffic rules, location semantics, weather conditions, and/or the level of fatigue of the driver. For situations involving more than one vehicle, pairs of maneuvers can be labeled as "conflicting" or "not conflicting" [75]. However, behavior-based risk estimation mostly focuses on driver behaviors, while it barely covers the behaviors of other traffic participants.

D. Challenges in Metric Design

There are some limitations of existing metrics. First, in most cases, if not all, the "metrics" are not strict mathematical metrics and don't satisfy all metric properties. As such, it is difficult to understand and analyze them across systems and scenarios in a commensurable fashion. Second, typically, even though the risk is measured on a continuous scale, for usability these measurements are thresholded (e.g., *Safe* vs *Hazardous*). This results in information-loss. Moreover, they also render the categorization subjective. Third, while the existing metrics can evaluate some types of risks, none of them provides a high-level explanation covering all driveability factors.

Even though not covered in this paper, it is also worth considering whether non-safety factors such as psychological and emotional factors of both the driver and the passengers should also be accounted for by the driveability metric. Ideally, autonomous driving should enable transportation that is not only safe but also enjoyable. To this end, a recent work [80] investigated ride comfort measures in autonomous cars, which are shown to be affected by factors such as motion sickness, naturality, apparent safety, etc. Some metrics for evaluation of driver's emotional feelings such as road frustration index has also been developed in industrial applications. However, it remains a challenging question how to harmonize the concept of driveability for both autonomous cars and the humans inside the vehicle.

As shown above, the actual scope of driveability metric depends greatly on the driving context and use case, which makes it challenging to design a single comprehensive and interpretable metric for driveability. Therefore, composite metrics should be thought of that include both basic measures that

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

are generalizable across systems and additional measures that can be tailored to the specific requirements of target application. Furthermore, such composite metrics could contain heterogeneous measures. Those can include but are not limited to continuous risk scores, the states and trajectories of all traffic participants [13], and even semantic descriptors such as "road construction" or "low lane marking visibility" which are shown to better communicate the intent of the autonomous car to its driver [81]. There is also dependency on the availability of up-to-date a-priori maps of the roads as well as cultural issues specific to a region. For example, driving in the US is probably easier than in some of the less developed countries where roads are less organized and people tend to disobey traffic rules more often.

IV. Datasets

The development of robust autonomous driving models depends on having access to large-scale training datasets, especially as more learning-based approaches are incorporated. Over the past decade, tens of datasets for autonomous driving have been collected and made public by multiple institutes around the world. These datasets are a valuable resource for the research community to develop benchmarks and consolidate research efforts. However, as autonomous driving encompasses numerous tasks in perception, localization and behavior analysis and the datasets are greatly varied in application focus, it is not trivial to determine which dataset to use for which task. Therefore, in this section, we provide an up-to-date comprehensive list of 54 publicly available datasets for autonomous driving and categorize them according to the tasks that they are suitable for. Particularly, since we have identified the factors that contribute to driveability and their associated challenges in Section II, we will highlight the datasets that can be used to address these challenges. We also publish this dataset archive at *https://sites.google.com/view/driveability-survey-datasets*, which allows interactive exploration of datasets and will continue to be maintained after publication.

Prior to this work, a survey of publicly available datasets was published in [131] and includes 27 publicly available datasets collected on public roads before late 2016. However, some of the largest and most diversified datasets have been released in the past two years, which are not included in [131]. We enrich the list in [131] by adding 31 more datasets most of which were released after 2016. We also exclude 4 datasets in [131] that have broken web links, impose a charge, or are integrated into newer datasets. We would like to mention that besides our selected datasets, there are other reported efforts on dataset acquisition such as the dataset "TorontoCity" collected by researchers at University of Toronto [132]. Unfortunately, this dataset no longer supports open access and is therefore not included in this paper.

Different from [131] which compares the datasets' metadata such as venue, volume, traffic condition, sensor setup and file type, we compare the datasets from an application perspective in terms of which tasks each dataset is suitable for, diversity and modality, level of annotations, and whether training/testing data and benchmarks are provided. Furthermore, we will summarize the trends emerging in these recently published datasets and propose directions for future dataset collection.

A. Overview of Datasets

We follow the same inclusion criteria to select relevant datasets as those used in [131], i.e., the dataset must be collected by on-board sensors of a vehicle running on public roads, contain camera or LiDAR data, and allow free open access. By an extensive search and snowballing in dataset websites, publications and competitions on autonomous driving, 54 datasets have been included in this survey and their metadata are presented in Table II and III. The symbol ($\star$) denotes that the dataset contains multiple subsets that are collected with different sensors or for different purposes. In the following, we elaborate on the aspects considered when presenting these datasets. While license information is not included in the dataset table, we suggest the readers check and observe the license when using these datasets.

Time & Venue denotes when the dataset was published and where the data was collected. In terms of time, half of the datasets were released since the year of 2016, which shows an increasing interest and effort on public dataset collection to boost the progress in autonomous driving research. In terms of venue, while most datasets published before 2016 were collected in Europe and the United States, the collection of many of the more recent datasets took place in Asian countries (such as ApolloScape, DBNet, IDD, KAIST, Road Damage Dataset, etc), and even around the globe (Mapillary Vistas and NEXET) with collaborative efforts from drivers worldwide who uploaded images to the database. Such diversity in locations is a leap forward towards enabling autonomous driving on road networks all over the world.

Data provided lists the data modalities and development codes provided by the dataset. A dataset is considered to contain videos if it provides either videos or image sequences that capture temporal information. Otherwise, a dataset is considered to only contain images if standalone images are provided without preceding or tailing video clips. For datasets that contain raw video clips with annotations made on selected but not all video frames, we mark them as containing both video and image data. Additional to videos and images, around one-fourth of the presented datasets provide LiDAR scans, GPS/IMU (inertial measurement unit) data, and/or vehicle status data including steering angle and velocity. There are also two datasets that provide infrared and thermal scans. Data collected from various types of sensors is greatly appreciated in deep multimodal learning, which is believed to enhance the inference performance of deep neural networks [133]. Most of the datasets provide codes in Python, Matlab or C++ for data preprocessing and visualization.

Annotation presents the type of labels provided by the dataset. The most common annotation types are bounding boxes and semantic labels, where the former is used for object detection and tracking and the latter is used for semantic segmentation. Some datasets also provide lane markings for lane detection. In addition to these graphical labels, a few datasets provide behavioral labels or data for higher-level

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

TABLE II

OVERVIEW OF PUBLICLY OPEN DATASETS FOR AUTONOMOUS DRIVING – PART I

Abbreviations used: VI-Video, IM-Image, Li-LiDAR, VD-Vehicle Data, CO-Codes; BB-Bounding Box, SL-Semantic Label, LM-Lane Marking, BL-Behavioral Label, O-Others; UR-Urban, RU-Rural, HI-Highway; WE-Weather, SE-Season, LO-Location, NI-Night, IL-Illumination; s-stereo images; uz-unzipped. Unit: K-Kilo, M-Million.

Dataset	Provider	Time & Venue	Data provided						Annotation				Traffic			Diversity				Volume
			VI	IM	Li	GPS	VD	CO	BB	SL	LM	BL	UR	RU	HI	WE	SE	NI	IL	
AMUSE [82]	Linköping Univ.	2013; Linköping, Sweden	✓			✓	✓	✓					✓			✓				7 VI, 1.2TB
Apollo Open Platform(⋆) [83]	Baidu Inc	2018; multiple cities in China	✓		✓		✓		✓				✓	✓	✓	✓				multiple datasets, volumes vary
ApolloScape [84]	Baidu Inc	2018; multiple cities in China	✓							✓	✓		✓	✓	✓	✓			✓	>140K IM
Argoverse(⋆) [85]	Argo AI	2019; Miami, Pittsburgh, US	✓	✓	✓	✓		✓	✓		✓		✓			✓	✓	✓	✓	2 datasets, 113 VI and >300K VI
Belgium Traffic Sign [86]	ETH Zürich	2011; Belgium		✓					✓				✓							>9K IM, 50GB
Berkeley DeepDrive [87]	UC Berkeley	2017; San Francisco Bay Area, New York, US	✓	✓		✓		✓	✓	✓	✓		✓	✓	✓	✓		✓	✓	100K IM, 1.8TB
Bosch Small Traffic Lights [88]	Bosch North America Research	2017; San Francisco Bay Area, US	✓						✓				✓	✓		✓			✓	>13K IM
Brain4Cars [89]	Cornell Univ.	2016; two states in US	✓			✓	✓	✓				✓	✓		✓				✓	700 VI
Caltech Pedestrian [90]	California Inst. of Tech.	2009; Los Angeles, US	✓					✓	✓				✓							250K IM, 11GB
CamVid [91]	Univ. of Cambridge	2009; Cambridge, UK	✓	✓				✓		✓			✓							700 IM, 8GB
CCSAD [92]	Centro de Investigacin en Matemticas	2014; Guanajuato, Mexico	✓			✓							✓					✓	✓	>96K IM(s), 500GB
CityScapes [93]	Daimler AG,MPI-IS,TU Darmstadt	2016; 50 cities in Germany, Switzerland & France	✓	✓		✓	✓	✓		✓			✓				✓			25K IM(s), 63GB
CMU [94]	Carnegie Mellon Univ.	2011; Pittsburgh, US	✓		✓	✓							✓			✓	✓		✓	16 VI, 275GB
Comma.ai [95]	comma.ai	2016; San Francisco, US	✓			✓	✓	✓							✓			✓		80GB
Comma2k19 [96]	comma.ai	2018; San Jose, US	✓			✓	✓	✓							✓			✓		2019 VI, 100GB
CULane [97]	Chinese Univ. of Hong Kong	2018; Beijing, China		✓				✓			✓		✓	✓	✓			✓	✓	133K IM
Daimler Pedestrian(⋆) [98]	Daimler AG R&D, Univ. of Amsterdam	2006-2016; Beijing, China, others unknown	✓	✓					✓	✓			✓							8 datasets, 2.5MB - 45GB each
DAVIS [99]	Univ. of Zürich, ETH Zürich	2017; Switzerland, Germany	✓			✓	✓	✓					✓	✓	✓	✓		✓	✓	41 VI, 450GB
DBNet [100]	Shanghai Jiao Tong Univ., Xiamen Univ.	2018; several cities in China	✓		✓		✓						✓	✓	✓					>10K IM, >1TB(uz)
DIPLECS(⋆) [101]	Univ. of Surrey	2015; Surrey, UK, Stockholm, Sweden	✓				✓					✓	✓	✓	✓					4.3GB & 1.1GB
Dr(eye)ve [102]	Univ. of Modena and Reggio Emilia	2016; Modena, Italy	✓				✓					✓	✓	✓	✓	✓		✓		74 VI, 555K IM
EISATS(⋆) [103]	Univ. of Auckland	2010; multiple locations in Germany and New Zealand	✓		✓							✓	✓	✓	✓	✓		✓	✓	multiple datasets, volumes vary
Elektra(⋆) [104]	Autonomous Univ. of Barcelona, Polytechnic Univ. of Catalonia	2016; Barcelona, Spain	VI & Infrared						✓				✓					✓		9 datasets, 0.5GB-12GB each
ETH Pedestrian [105]	ETH Zürich	2009; Zürich, Switzerland	✓						✓				✓							>4.8K IM(s), 660MB
FLIR [106]	FLIR Systems Inc	2018; Santa Barbara, US	VI, IM, Infrared						✓				✓		✓	✓	✓	✓	✓	>14K IM
Ford [107]	Univ. of Michigan	2009; Michigan, US	✓		✓	✓		✓					✓							80GB & 120GB
German Traffic Sign [108]	Ruhr Univ. Bochum	2012; Germany		✓					✓											5K IM, 1.6GB
HCI Challenging Stereo [109]	HCI(Heidelberg), Bosch Corporation Research	2012; Hildesheim, Germany	✓										✓			✓		✓	✓	11 VI(s)
HD1K [110]	HCI(Heidelberg), Robert Bosch GmbH	2018; Heidelberg, Germany	✓						optical flow				✓			✓		✓	✓	>1K IM
Highway Workzones [37]	Carnegie Mellon Univ.	2015; US	✓						✓						✓	✓	✓		✓	6 VI, 1.2GB
IDD [111]	IIIT Hyderabad, Intel Bangalore	2019; Hyderabad, Bangalore cities, India	✓	✓				✓		✓			✓	✓		✓		✓	✓	10K IM
JAAD [66]	York University	2016; mostly in Ukraine and Canada	✓						✓			✓	✓			✓		✓		347 VI, 170GB
KAIST Multi-Spectral [112]	KAIST	2015; South Korea	VI, Li, GPS, CO, Infrared						✓				✓					✓	✓	10 VI
KAIST Urban [113]	KAIST	2018; multiple cities in South Korea	✓		✓	✓		✓					✓							19 VI, 1GB-22GB each
KITTI(⋆) [114]	Karlsruhe Inst of Technology, Toyota Technological Inst	2011; Karlsruhe, Germany	✓	✓	✓	✓		✓	✓	✓	✓		✓	✓	✓				✓	multiple datasets, volumes vary

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

TABLE III

OVERVIEW OF PUBLICLY OPEN DATASETS FOR AUTONOMOUS DRIVING – PART II

Abbreviations used: VI-Video, IM-Image, Li-LiDAR, VD-Vehicle Data, CO-Codes; BB-Bounding Box, SL-Semantic Label, LM-Lane Marking, BL-Behavioral Label, O-Others; UR-Urban, RU-Rural, HI-Highway; WE-Weather, SE-Season, LO-Location, NI-Night, IL-Illumination; s-stereo images; uz-unzipped. Unit: K-Kilo, M-Million.

Dataset	Provider	Time & Venue	Data provided						Annotation				Traffic			Diversity				Volume
			VI	IM	Li	GPS	VD	CO	BB	SL	LM	BL	UR	RU	HI	WE	SE	NI	IL	
LISA Traffic Sign [115]	Univ. of California, San Diego	2012; US		✓				✓	✓											6.6K IM, 8GB
LostAndFound [48]	Daimler AG	2016, Germany		✓			✓			✓			✓							2K IM(s), 40GB
Málaga [116]	Univ. of Málaga	2014; Málaga, Spain	✓		✓	✓		✓					✓							15 VI(s), 70GB
Mapillary Vistas [117]	Mapillary AB	2017; around the globe		✓						✓			✓	✓		✓	✓	✓	✓	25K IM
NEXET [118]	Nexar	2017; around the globe		✓				✓	✓				✓	✓	✓	✓	✓	✓	✓	55K IM, 10GB
NightOwls [63]	Oxford Univ., Max Planck Inst for Informatics, Continental Corp	2018; 7 cities, Germany, Netherlands & UK	✓	✓					✓				✓			✓	✓	✓	✓	279K IM
nuScenes [119]	nuTonomy Inc, Aptiv	2018; Boston, US & Singapore	✓	✓	✓	✓		✓	✓			✓	✓	✓		✓	✓	✓	✓	1K VI, 1.4M IM
Oxford RobotCar [120]	Oxford Univ.	2015; Central Oxford, UK	✓		✓	✓	✓	✓					✓			✓	✓		✓	130 VI(s), 23TB
Road Damage [35]	Univ. of Tokyo	2018; multiple cities in Japan		✓				✓	✓				✓			✓				9K IM
Stanford Track [121]	Stanford Univ.	2010; Stanford Univ., US			✓	✓		✓		stixel			✓							14K tracks, 5.7GB
Stixel [122]	Daimler AG	2013; Germany	✓				✓			stixel					✓	✓				2.5K IM, 3GB
StreetLearn [123]	DeepMind	2019; New York City & Pittsburgh, US		✓		✓		✓					✓							114K IM
Sydney Urban [124]	Univ. of Sydney	2013; Sydney, Australia			✓			✓					✓							200MB
TME Motorway [125]	Czech Technical Univ.	2011; Northern Italy	✓					✓	✓						✓				✓	28 VI
TUD-Brussels Pedestrian [126]	Max Planck Inst for Informatics	2009; Belgium	✓						✓				✓							1.6K IM
TuSimple(⋆) [127]	TuSimple	2017; venue unknown	✓	✓				✓	✓		✓				✓	✓				7K IM & 5K IM
UAH [128]	Univ. of Alcalá	2016; Madrid, Spain	✓			✓	✓	✓				✓			✓					35 VI
Udacity(⋆) [129]	Udacity	2016; Mountain View, US	✓		✓	✓	✓	✓	✓				✓						✓	300GB(uz)
UTBM Robocar [130]	Univ.of Tech. of Belfort-Montbéliard	2018; Montbéliard, France	✓		✓	✓	✓	✓					✓			✓	✓	✓		11 VI(s)

(a) bounding box (b) semantic label (c) lane marking (d) driveable area

(e) driver's gaze (f) stixel label (g) depth image (h) optical flow

Fig. 2. Sample images with various annotation types. (a)(c)(d) are from Berkeley DeepDrive, (b) is from CityScapes, (e) is from Dr(eye)ve, (f) is from Stixel, (g) is from ApolloScape, and (h) is from HD1K. All sample images are extracted from the webpages hosting the datasets.

scene reasoning. Such labels include synchronized videos of both the driver's face and the road (Brain4Cars), driver's gaze (Dr(eye)ve, Elektra) and behaviors of both the drivers and pedestrians present in the same scene (JAAD). Besides these annotations, the optical flow information is provided by the HD1K dataset, and a stixel label that uses multiple cubics to represent an object is provided in the Stanford Track and the Stixel dataset. Figure 2 shows one example for each type of annotation. Data labeling is generally considered to be time-consuming and cumbersome, and therefore choosing a dataset with annotations could ease the labeling burden in model training.

Traffic records the traffic conditions where the majority of datasets focus on urban traffic. *Diversity* shows the diversity in environmental conditions under which the dataset was collected, which includes *weather, season, night and illuminations*. Two-thirds of the presented datasets provide diversified data in at least one of the above categories. Seasonal changes are not commonly seen due to the fact that most datasets are collected over shorter durations.

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
DATASET CATEGORIZATION BY AUTONOMOUS DRIVING TASKS

✓denotes that separate training and test sets are provided; ★ denotes that benchmark results are provided.

Task	Datasets
Stereo / 3D vision	CamVid, CCSAD, CMU, EISATS, Elektra, HCI Challenging Stereo, KAIST Urban, KITTI(✓★), Málaga, Oxford Robotcar, Stixel
Optical flow	HCI Challenging Stereo, HD1K, KITTI(✓★)
Object detection	Multi-class: Apollo Open Platform, Argoverse(✓), Berkeley DeepDrive(✓★), CamVid, CityScapes(✓), Elektra, FLIR, JAAD, KAIST Multi-Spectral, KAIST Urban, KITTI(✓★), NEXET(✓), nuScenes, Stanford Track, Sydney Urban, TME Motorway, Udacity(✓) Traffic sign (TS): Belgium TS(✓), Bosch Small Traffic Lights(✓★), Highway Workzones, German TS(✓★), LISA TS Pedestrian (Ped): Caltech Ped(✓★), Daimler Ped(✓★), ETH Ped, NightOwls(✓★), TUD-Brussels Ped(✓) Obstacle: LostAndFound(✓★), Road Damage(✓★)
Object tracking	Apollo Open Platform, Argoverse(✓★), Berkeley DeepDrive(✓), Caltech Ped(✓), Daimler Ped, Elektra, ETH Ped, FLIR, JAAD, KAIST Multi-Spectral, KITTI(✓★), NightOwls, nuScenes, Stanford Track, TME Motorway, TUD-Brussels Ped(✓), TuSimple(✓), Udacity(✓)
Lane/Road detection	ApolloScape(✓), Berkeley DeepDrive(✓), CULane(✓★), KAIST Multi-Spectral(★), KITTI(✓★), TuSimple(✓★)
Semantic segmentation	ApolloScape(✓★), Berkeley DeepDrive(✓), CamVid, CityScapes(✓★), Daimler Ped(✓), Elektra, IDD(✓★), KITTI(✓★), Mapillary Vistas(✓★)
Localization / SLAM	AMUSE, Argoverse(✓), CMU, Comma2k19, Ford, KAIST Multi-Spectral, KAIST Urban, KITTI(✓★), Málaga, Oxford Robotcar, StreetLearn(★), Udacity(✓★), UTBM Robocar(★)
End-to-end learning	Apollo Open Platform, Comma.ai, Comma2k19, DAVIS, DBNet(★), DIPLECS, Udacity(✓★)
Behavior analysis	Brain4Cars(★), DIPLECS, Dr(eye)ve, EISATS, Elektra, JAAD, UAH

Volume shows the total dataset size in terms of zipped files if not specified otherwise, which provides reference for disk usage when one considers storing and utilizing the dataset.

B. Dataset Categorization

As presented in Table II and III, nearly no two datasets are exactly the same in terms of what they offer. The datasets are usually collected and annotated for different purposes, and therefore, they are suitable for different autonomous driving tasks. To provide a practical view of how to use these datasets, in Table IV, we categorize the datasets according to the tasks shown in Fig. 1. Even though for a specific task, we only include datasets that were originally collected to perform the task, it should be noted that other datasets may also be applicable with additional labeling or data processing. We also show whether a dataset includes separate training and test sets, and whether the providers offer benchmark results online or in publications, which are very useful for comparison of methods developed using the dataset.

From Table IV, it can be observed that most of the datasets were collected for object detection, which is one of the early research fields for autonomous driving. There are also a good collection of datasets for object tracking and semantic segmentation. However, only few datasets can be used for reasoning about a scene at a higher level using behavior analysis, which is a research topic that has been studied insufficiently. Furthermore, as the end-to-end approach has been shown to have full potential of driving policy learning [134], a couple of datasets were published recently to encourage research in this direction.

C. Dataset Highlights for Driveability Assessment

While classical datasets such as KITTI provide canonical benchmarks for testing and comparison of baseline algorithms, datasets that capture more adverse scenes or provide different levels of annotations will be more interesting for improving the robustness of autonomous cars. The collection of challenging scenes helps to identify the limitations of state-of-the-art approaches and inspire novel and more robust algorithms. Fortunately, towards this goal, the following trends have been observed in recently released datasets, namely, high complexity and diversity, capturing potentially hazardous events, and providing behavioral and contextual data for prediction and inference. In the following, we highlight the datasets according to these three characteristics and take a deeper dive into what is offered in each dataset. The highlighted datasets are also associated with driveability factors in Table I.

1) High Diversity and Complexity: To capture realistic driving scenes, the data collected should be as natural and diversified as possible in terms of weather, traffic condition, illumination, etc. The most diverse datasets are Mapillary Vistas and NEXET, which consist of images uploaded from drivers all over the globe covering six continents. NEXET was collected by using the Nexar's dashcam and almost half of the images were taken at night time [118]. Bounding boxes for five classes of vehicles are provided. Compared to NEXET, Mapillary Vistas is also diverse in terms of image sources as the images were taken from different devices including both cameras and mobile phones and by photographers of varying experiences. However, only images with at least 1920×1080 resolution were selected [117]. Semantic labels for 66 classes of objects are provided by Mapillary Vistas in the research edition, whereas the commercial edition provides labels for 100 classes.

Another two large datasets for semantic segmentation are ApolloScape and CityScapes. CityScapes has gained significant popularity since its release. It provides 5K stereo images with pixel-level semantic labels and 20K stereo images with instance-level semantic labels [93]. The annotated image is taken from a 1.8-second video clip and the raw video clips are provided. 30 classes of objects are annotated and additional information such as outside temperature and precomputed disparity depth maps are also provided. However, the limitation of CityScapes is that most images were collected in daytime under good to medium weather condition. ApolloScape is by far the largest dataset with semantic labels covering 25 classes. Additionally, it provides 28 lane marking classes and depth images. Compared to CityScapes, ApolloScape captures more scenes with bad weather, reflections on vehicles and extreme lighting conditions [84]. The largest dataset that provides the most sensor measurements is nuScenes, which contains 1000 20-second long videos with LiDAR, Radar, camera, IMU and

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

TABLE V
NUMBER OF INSTANCES IN TRAINING AND VALIDATION SETS

Dataset	Total ($\times 10^3$)		Average per image	
	Person	Vehicle	Person	Vehicle
ApolloScape	543	1989	1.1/6.2/16.9	12.7/24.0/38.1
CityScapes	24.4	41	7.0	11.8
KITTI	6.1	30.3	0.8	4.1
Berkeley DeepDrive	86	-	1.2	-
Caltech	192	-	1.5	-

GPS data. It also provides 3D bounding boxes over 25 classes of objects annotated at 2Hz.

Berkeley DeepDrive provides the most comprehensive annotations with bounding boxes for 10 classes and lane markings for 8 classes on 100K images, and semantic labels for 40 classes on 10K images. The raw videos of 40 seconds are provided where the annotated images are extracted. Additionally, a unique annotation of driveable areas and a tag that shows the weather, time of day and scene context (residential, highway, tunnel, etc) of each image are also provided [87].

While most datasets were collected during a short time span, the CMU, Oxford RobotCar and UTBM Robocar datasets were collected over months traversing the same routes multiple times, which capture long-term changes in the environment. Particularly, Oxford RobotCar was collected on the same route twice a week over a year and provides various types of data including stereo and monocular images, 2D and 3D LiDAR scans, GPS, and inertial sensor data, which is good for research on long-term simultaneous localization and mapping (SLAM) in dynamic urban environments [120].

Scenes recorded in cluttered urban environments with more traffic participants are also favorable to enhance the robustness of learning models to deal with complex scenarios. Table V provides a comparison of the number of instances labeled in some large datasets which make this data available. ApolloScape is divided into three complexity levels with different numbers of instances per image [84]. ApolloScape and CityScapes contain images with more people as most of the data was collected in urban areas, whereas KITTI and Berkeley DeepDrive contain many scenes on highways with fewer people.

2) Hazardous Events: As identified in Section II, hazardous factors include road damages, construction, adverse weathers, hard-to-identify obstacles due to size or occlusion, and so on. Several datasets have been collected to address these challenges. In terms of road condition, Road Damage is the only dataset that annotates road surface damages. The images were taken by a vehicle-mounted smartphone and 8 types of road damages were identified according to the Road Maintenance and Repair Guidebook in Japan including liner crack, alligator crack, bump, line blur, etc [35]. CCSAD was also collected in locations with hazardous road conditions including irregular speed humps, abundant potholes, and peculiar flows of pedestrians [92]. However, only raw data is provided in CCSAD. IDD is the first large-scale dataset collected in India that captures many local complex driving conditions, which can be used to train safer autonomous driving algorithms that adapt to more crowded and unstructured environments. For road construction, Highway Workzones annotates signs specifically related to workzones, which is, however, limited to only highway traffic.

Among the many datasets that provide multi-class labels for object detection, LostAndFound is a unique dataset that annotates small obstacles caused by lost cargos down to the height of 5 centimeters [48]. In total, 42 classes of small obstacles are labeled such as crate, cardboard box, plastic bag, ball, and many others. In a similar vein, Bosch Small Traffic Lights provides videos that contain traffic lights with a width as small as 2 pixels, which are even difficult for human eyes to distinguish [88]. In total, 13 types of traffic lights including their shapes and colors are annotated.

Regarding weather conditions, many datasets contain scenes with rain, overcast, or even snow. Particularly, HCI Challenging Stereo provides 11 challenging sequences taken under extreme weather conditions which include scenes with flying snow, rain flares at night, snow at night, rain blurs, etc [109].

3) Behavioral and Contextual Data: For the purpose of driver attention analysis, driver's gaze fixation is usually collected. While both DIPLECS and Elektra contain a moderate-sized data subset that provides the gaze information, the only large-scale dataset with accurately measured gaze fixation is Dr(eye)ve. In Dr(eye)ve, the gaze data is projected onto the video captured by a front camera, which can be used to predict driver's intent, improve road safety, and plan better driving strategies [102]. Instead of gaze data, Brain4Cars recorded videos of drivers inside the car. Together with the videos of the road and the labels of six classes of vehicle maneuvers such as stop, turn and lane change, Brain4Cars is used for anticipating maneuvers several seconds in advance [89]. Another dataset that focuses on driver's behavior is UAH, which simulates normal, drowsy and aggressive driving and records 7 maneuvers such as lane-drifting, overspeeding, car-following, etc.

While the above datasets only focus on driver behavior analysis, JAAD labels the behaviors of both drivers and pedestrians occurring in the same scene. Driver behaviors include stopping, moving slow, accelerating, etc, whereas pedestrian behaviors include actions such as crossing, looking, slowing down as well as their moving directions and whether they are at an intersection [66]. The JAAD dataset is very useful for studying joint attention of drivers and pedestrians, which is still a challenging problem that awaits effective solutions [67], as mentioned in Section II-C.

V. FILLING THE DATA GAP

While there is an increased emphasis on dataset acquisition, the available data is still insufficient for learning an autonomous driving model that can operate robustly anywhere at anytime. In this section, we will discuss different ways to fill this data gap. We address the following three questions: 1) in which areas do we need to collect more naturalistic driving datasets? 2) how can one utilize synthetic data to complement real data? and 3) how can one use driving simulators for developing and testing learning algorithms? We conduct a brief discussion on current research efforts, open issues, and possible future directions related to these topics.

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

A. Targeted Data Acquisition

The robustness of autonomous driving models depends on continuously resolving real-world hazardous corner cases, which requires continuously collecting datasets that expose the true diversity of the driving environment globally and include rich scene representations. In the following, we list several directions for new dataset collection that complement existing datasets.

- For joint attention studies and intent prediction, datasets that simultaneously record the behaviors of pedestrians, cyclists, and drivers of both the ego car and neighboring vehicles are needed. Behaviors include but are not limited to facial expressions, gestures and gaze. It would be better that such data is collected during an interactive event between multiple road users, e.g., a pedestrian establishes eye contact with the driver or the car before crossing the street, because many accidents can be avoided by interactions using eye contact or simple gestures [67].
- For robust obstacle detection, more datasets that capture uncommon hazards are needed, such as small pieces of cargo or debris falling from a vehicle in front of the ego car, accident scenes, road construction, etc. Recently, a Near-miss Incident DataBase was introduced in [18], but it is now under reconstruction and charges a fee for access.
- For deep scene understanding, datasets with multi-level annotations are helpful which include low-level object annotation, mid-level trajectory annotation and high-level behavior and relationship annotation [52].
- For driveability evaluation of various types of scenes, contextual scene descriptions are needed. The data collected is better tagged according to driving contexts such as intersections/roundabouts/non-intersections, signaled/unsignaled pedestrian crossing, number of marked/unmarked lanes, and so on, which could facilitate adaptive driving policy learning in various driving contexts.

Though not discussed here, the data collected could have dependency on the hardware and sensor calibration used during data acquisition. Many datasets provide calibration files for users to get exact positions of sensors and better use the measurements. Nonetheless, the goal is to learn an autonomous driving model that is invariant to hardware changes and work across different platforms.

B. Synthesizing Data

It can be costly and cumbersome to collect data in the physical world and annotate them at a large scale. Therefore, researchers also resort to synthetic datasets that contain visually realistic images and automatically generated annotations for autonomous driving studies. Generating synthetic data is considered to be one major data augmentation approach for learning-based algorithms [135]. Synthesizing data can create additional samples capturing conditions not covered in the real-world datasets, and can be used to augment both training and testing sets. Some well established synthetic driving datasets include Virtual KITTI [136], Synthia [137], VIPER [138], etc, which are generally produced by video game engines.

A big concern with synthetic data is that it is dependent on the data generation model and therefore may present bias and not generalize well enough to the physical world. Fortunately, thanks to the continuous development of game engines and learning algorithms such as generative adversarial networks (GANs) [139], the realism of synthetic images keeps increasing, which can then better generalize the models trained on these synthetic data to real-world scenarios. For example, it has been shown in [140] that a CNN-based hand pose estimator trained only on synthetic data generated by GAN can even outperform a CNN trained on real images when applied to a realistic test dataset. Similar approaches are worth pursuing for training autonomous driving models.

Besides training, synthetic data can also be used to test and identify the limitations of the trained model by simulating not yet encountered conditions. To this end, two very recent studies [1], [62] focus on DNN testing by generating realworld-like images that cover scenes with extreme weather conditions to detect erroneous behaviors of a trained deep end-to-end learning model, where [1] uses affine image transformations and [62] uses GAN to generate the test images. The underlying assumption is that at the same location, the predicted car maneuver should be similar under different weather conditions or illuminations. Both approaches have been reported to detect thousands of erroneous behaviors of top-ranked DNN models in the Udacity challenge [129]. While these findings are enlightening, the coverage of the generated test images in both studies are still limited to only rain and snow conditions. Extended synthetic datasets are needed to test DNNs in additional conditions and in more complex scenarios. Further robust mechanisms to assess and verify generalizability to real-world are needed.

C. Driving Simulators

As pointed out in [141], it would need test driving for hundreds of millions of miles without accident to prove that an autonomous driving system is safe enough to be adopted. This is in general impractical and still cannot cover all conceivable scenarios. Therefore, simulators play an important role in simulating different scenarios and performing exhaustive tests for both ADAS and autonomous cars. Many simulators have been developed in recent years, such as CARLA [142], Microsoft's AirSim [143], Baidu's Apollo [144], and NVIDIA's DRIVE Constellation [145], just to name a few. These simulators generate driving scenarios from user-defined models of roads, traffic and vehicles, which allow testing of autonomous driving models and discovery of interrelationship among various driveability factors. However, even though high-fidelity environments are simulated, the conditions there are generally fully controlled and can be different from real-world situations that have higher uncertainty and dynamics. Therefore, one should be aware that the conclusions drawn from using these simulators may not be easily translated to real-world scenarios.

With various data sources of both realistic and simulated data, a collective approach is needed to integrate them and reap the most benefit from all available data sources. To this

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

end, the PEGASUS research project [146] develops a data processing pipeline that stores data from different sources such as field tests, driving simulator studies, and traffic simulations into a database [147]. The main objective of constructing this database is to collect relevant traffic scenarios and establish a common evaluation basis for autonomous vehicle testing and validation.

VI. Further Discussion

While this paper focuses on supervised learning that requires sufficient data, we would like to point out that additional advancements are needed in artificial intelligence (AI) to learn from limited data and generalize to unknown environments. Ideally, the AI engine that powers autonomous driving should have the capability of transferring models trained on source data to any target domain. For example, if a driving model is trained under good lighting condition during midday, will it have similar performance at all times during the day, at different geographic locations, and under all weather conditions? Building a robust model that performs well in all those conditions may require extension or augmentation through unsupervised learning and reinforcement learning techniques that can transfer knowledge and driving policy to unseen scenarios.

Many reinforcement learning based approaches have been reported for driving policy adaptation. For example, a RNN-based approach is proposed in [148] which first learns an optimal initial policy architecture from expert demonstration, and then adapts this policy to a new driving domain using the rewards obtained in the new domain. To better understand a new driving context, it is also important to find a representation of a scene such that the difference between the source and the target scenes and the error made by learning models can be minimized [149]. To address this issue, [150] adds semantic clues from the environment such as the distance of a pedestrian from the curbside and traffic light status to make pedestrian motion prediction more robust and flexible in new environments. Overall, the questions of what to transfer, when to transfer and how to transfer [149] need to be understood better in the context of autonomous driving.

One last remark we would like to make here is that robust AI is only one component contributing to safe autonomous driving. Achieving a safe autonomous vehicle requires solving interdisciplinary problems in domains of computing hardware, robotics, security, social acceptance and many others [68], that are beyond the scope of this paper.

VII. Concluding Remarks

In this paper, we have reviewed recent research efforts from a driveability perspective for autonomous driving. We presented both explicit and implicit factors that contribute to scene driveability, identified the potential risks posed by these factors, and investigated existing methods and metrics used for driveability assessment. With these investigations, we have shown the necessity of pursuing principled metrics for driveability, which can be represented by either a set of novel sophisticated metrics or a composition of metrics. It should also be well understood how the driveability metric interacts with other metrics used in system level verification and validation, which will have implication on optimization of multiple metrics concurrently and the trade-off therein.

Furthermore, we have conducted an exhaustive overview of 54 open datasets collected on public roads and categorized these datasets according to use cases. More importantly, we highlighted datasets that are more suitable for training robust driving models and identified the scenarios that need more data acquisition. We have also proposed ways to fill the data gap including conducting targeted dataset acquisition, using synthetic data for training and testing, exploring driving simulators, and transferring knowledge to unseen scenarios. We hope this paper can serve both as a guidance on dataset selection and construction and as an invitation to pursue novel approaches that enable autonomous cars to navigate through all environments safely and reliably.

Acknowledgement

The authors would like to thank Prof. Maxim Likhachev from Carnegie Mellon University for his invaluable comments that improved the manuscript.

References

[1] Y. Tian, K. Pei, S. Jana, and B. Ray, "DeepTest: Automated testing of deep-neural-network-driven autonomous cars," Aug. 2017, *arXiv:1708.08559*. [Online]. Available: https://arxiv.org/abs/1708.08559

[2] D. Muoio. *6 Scenarios Self-Driving Cars Still Can't Handle*. Accessed: Oct. 15, 2018. [Online]. Available: https://tinyurl.com/y8kgnxoo

[3] S. Ulbrich, T. Menzel, A. Reschka, F. Schuldt, and M. Maurer, "Defining and substantiating the terms scene, situation, and scenario for automated driving," in *Proc. IEEE 18th Int. Conf. ntell. Transp. Syst. (ITSC)*, Sep. 2015, pp. 982–988.

[4] S. Hecker, D. Dai, and L. Van Gool, "Failure prediction for autonomous driving," May 2018, *arXiv:1805.01811*. [Online]. Available: https://arxiv.org/abs/1805.01811

[5] S. Kumar, S. Gollakota, and D. Katabi, "A cloud-assisted design for autonomous driving," in *Proc. 1st Ed. MCC Workshop Mobile Cloud Comput.*, 2012, pp. 41–46.

[6] S. Cafiso, G. Cava, and A. Montella, "Safety inspections as supporting tool for safety management of low-volume roads, " *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2203, no. 1, pp. 116–125, 2011.

[7] G. E. Fuchs, "Risk based automotive insurance rating system," U.S. Patent 0 189 303 A1, Jun. 30, 2016.

[8] C. Hsu-Hoffman, R. Madigan, and T. McKenna, "Subjective route risk mapping and mitigation," U.S. Patent 0 221 149 A1, Aug. 3, 2017.

[9] Y. Tsukada, T. Okutani, S. Itsubo, and J. Tanabe, "Evaluation of roads network in japan from viewpoint of drivability," in *Proc. Eastern Asia Soc. Transp. Stud.*, 2007, p. 336.

[10] E. Bekiaris, A. Amditis, and M. Panou, "Drivability: A new concept for modelling driving performance," *Cognition, Technol. Work*, vol. 5, no. 2, pp. 152–161, 2003.

[11] J. Leonard *et al.*, "A perception-driven autonomous urban vehicle," *J. Field Robot.*, vol. 25, no. 10, pp. 727–774, 2008.

[12] S. Thrun *et al.*, "Stanley: The robot that won the DARPA grand challenge," *J. Field Robot.*, vol. 23, no. 9, pp. 661–692, 2006.

[13] S. Sivaraman and M. M. Trivedi, "Dynamic probabilistic drivability maps for lane change and merge driver assistance," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2063–2073, Oct. 2014.

[14] K. Kim, K. Lee, B. Ko, and K. Yi, "Design of integrated risk management-based dynamic driving control of automated vehicles," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 1, pp. 57–73, Jan. 2017.

[15] D. I. F. Ferguson, A. Wendel, Z. Xu, D. H. Silver, and B. D. Luders, "Determining drivability of objects for autonomous vehicles," U.S. Patent 10 203 696 B2, Feb. 12, 2012.

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

[16] O. Scheel, L. Schwarz, N. Navab, and F. Tombari, "Situation assessment for planning lane changes: Combining recurrent models and prediction," May 2018, *arXiv:1805.06776*. [Online]. Available: https://arxiv.org/abs/1805.06776

[17] I. Chowdhury, "A user-centered approach to road design: Blending distributed situation awareness with self-explaining roads," Ph.D. dissertation, Inst. Infrastruct. Environ., Heriot-Watt Univ., Dubai, UAE, 2014.

[18] R. Takahashi, N. Inoue, Y. Kuriya, S. Kobayashi, and K. Inui, "Explaining potential risks in traffic scenes by combining logical inference and physical simulation," *Int. J. Mach. Learn. Comput.*, vol. 6, no. 5, pp. 248–255, 2016.

[19] S. Singh, "Critical reasons for crashes investigated in the national motor vehicle crash causation survey," Traffic Saf. Facts-Crash Stats, Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep. HS 812 506, Mar. 2018,

[20] P. Bhavsar, P. Das, M. Paugh, K. Dey, and M. Chowdhury, "Risk analysis of autonomous vehicles in mixed traffic streams," *Transp. Res. Rec.*, vol. 2625, no. 1, pp. 51–61, 2017.

[21] *International Road Assessment Programme*. Accessed: Jan. 10, 2019. [Online]. Available: https://www.irap.org/

[22] A. Gern, R. Moebus, and U. Franke, "Vision-based lane recognition under adverse weather conditions using optical flow," in *Proc. IEEE Intell. Vehicle Symp.*, vol. 2, Jun. 2002, pp. 652–657.

[23] Z. Wang, G. Cheng, and J. Y. Zheng, "All weather road edge identification based on driving video mining," in *Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2017, pp. 1–6.

[24] U. Lee, J. Jung, S. Jung, and D. H. Shim, "Development of a self-driving car that can handle the adverse weather," *Int. J. Automot. Technol.*, vol. 19, no. 1, pp. 191–197, 2018.

[25] J. M. Alvarez, A. Lopez, and R. Baldrich, "Illuminant-invariant model-based road segmentation," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2008, pp. 1175–1180.

[26] P. Corke, R. Paul, W. Churchill, and P. Newman, "Dealing with shadows: Capturing intrinsic scene appearance for image-based outdoor localisation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2013, pp. 2085–2092.

[27] W. Maddern, A. Stewart, C. McManus, B. Upcroft, W. Churchill, and P. Newman, "Illumination invariant imaging: Applications in robust vision-based localisation, mapping and classification for autonomous vehicles," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 1–8.

[28] Y.-L. Chen, Y.-H. Chen, C.-J. Chen, and B.-F. Wu, "Nighttime vehicle detection for driver assistance and autonomous vehicles," in *Proc. IEEE 18th Int. Conf. Pattern Recognit. (ICPR)*, vol. 1, Aug. 2006, pp. 687–690.

[29] H. Kuang, L. Chen, F. Gu, J. Chen, L. Chan, and H. Yan, "Combining region-of-interest extraction and image enhancement for nighttime vehicle detection," *IEEE Intell. Syst.*, vol. 31, no. 3, pp. 57–65, May 2016.

[30] J. Wagner, V. Fischer, M. Herman, and S. Behnke, "Multispectral pedestrian detection using deep fusion convolutional neural networks," in *Proc. 24th Eur. Symp. Artif. Neural Netw., Comput. Intell. Mach. Learn.*, 2016, pp. 509–514.

[31] A. González *et al.*, "Pedestrian detection at day/night time with visible and FIR cameras: A comparison," *Sensors*, vol. 16, no. 6, p. 820, 2016.

[32] M. S. Shirazi and B. T. Morris, "Looking at intersections: A survey of intersection monitoring, behavior and safety analysis of recent studies," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 4–24, Jan. 2017.

[33] A. Geiger, M. Lauer, C. Wojek, C. Stiller, and R. Urtasun, "3D traffic scene understanding from movable platforms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 5, pp. 1012–1025, May 2014.

[34] B. Okumura *et al.*, "Challenges in perception and decision making for intelligent automotive vehicles: A case study," *IEEE Trans. Intell. Veh.*, vol. 1, no. 1, pp. 20–32, Mar. 2016.

[35] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection using deep neural networks with images captured through a smartphone," Jan. 2018, *arXiv:1801.09454*. [Online]. Available: https://arxiv.org/abs/1801.09454

[36] P. Papadakis, "Terrain traversability analysis methods for unmanned ground vehicles: A survey," *Eng. Appl. Artif. Intell.*, vol. 26, no. 4, pp. 1373–1385, 2013.

[37] Y.-W. Seo, J. Lee, W. Zhang, and D. Wettergreen, "Recognition of highway workzones for reliable autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 708–718, Apr. 2015. [Online]. Available: http://www.andrew.cmu.edu/user/jonghole/workzone/data/

[38] A. B. Hillel, R. Lerner, D. Levi, and G. Raz, "Recent progress in road and lane detection: A survey," *Mach. Vis. Appl.*, vol. 25, no. 3, pp. 727–745, 2014.

[39] R. Danescu and S. Nedevschi, " Probabilistic lane tracking in difficult road scenarios using stereovision," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 2, pp. 272–282, Feb. 2009.

[40] Q. Li, L. Chen, M. Li, S.-L. Shaw, and A. Nüchter, "A sensor-fusion drivable-region and lane-detection system for autonomous vehicle navigation in challenging road scenarios," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 540–555, Feb. 2014.

[41] J. Bao, Y. Zhang, X. Su, and R. Zheng, "Unpaved road detection based on spatial fuzzy clustering algorithm," *EURASIP J. Image Video Process.*, vol. 2018, no. 1, p. 26, 2018.

[42] T. Suzuki, H. Kataoka, Y. Aoki, and Y. Satoh, "Anticipating traffic accidents with adaptive loss and large-scale incident DB," Apr. 2018, *arXiv:1804.02675*. [Online]. Available: https://arxiv.org/abs/1804.02675

[43] J. Janai, F. Güney, A. Behl, and A. Geiger, "Computer vision for autonomous vehicles: Problems, datasets and state-of-the-art," Apr. 2017, *arXiv:1704.05519*. [Online]. Available: https://arxiv.org/abs/1704.05519

[44] S. Sivaraman and M. M. Trivedi, "Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1773–1795, Dec. 2013.

[45] A. Mukhtar, L. Xia, and T. B. Tang, "Vehicle detection techniques for collision avoidance systems: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2318–2338, Oct. 2015.

[46] D. Geronimo, A. M. Lopez, A. D. Sappa, and T. Graf, "Survey of pedestrian detection for advanced driver assistance systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1239–1258, Jul. 2010.

[47] X. Li *et al.*, "A new benchmark for vision-based cyclist detection," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2016, pp. 1028–1033.

[48] P. Pinggera, S. Ramos, S. Gehrig, U. Franke, C. Rother, and R. Mester, "Lost and found: Detecting small road hazards for self-driving vehicles," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 1099–1106. [Online]. Available: http://www.6d-vision.com/lostandfounddataset

[49] C. Zhou, J. Yang, C. Zhao, and G. Hua, "Fast, accurate thin-structure obstacle detection for autonomous mobile robots," Aug. 2017, *arXiv:1708.04006*. [Online]. Available: https://arxiv.org/abs/1708.04006

[50] E. Ohn-Bar and M. M. Trivedi, "Are all objects equal? Deep spatio-temporal importance prediction in driving videos," *Pattern Recognit.*, vol. 64, pp. 425–436, Apr. 2017.

[51] D. Bevly *et al.*, "Lane change and merge maneuvers for connected and automated vehicles: A survey," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 1, pp. 105–120, Mar. 2016.

[52] J.-R. Xue, J.-W. Fang, and P. Zhang, "A survey of scene understanding by event reasoning in autonomous driving," *Int. J. Autom. Comput.*, vol. 15, no. 3, pp. 249–266, 2018.

[53] S. Dixit *et al.*, "Trajectory planning and tracking for autonomous overtaking: State-of-the-art and future prospects," *Annu. Rev. Control*, vol. 45, pp. 76–86, Jan. 2018.

[54] X. Geng, H. Liang, B. Yu, P. Zhao, L. He, and R. Huang, "A scenario-adaptive driving behavior prediction approach to urban autonomous driving," *Appl. Sci.*, vol. 7, no. 4, p. 426, 2017.

[55] J. Schulz, C. Hubmann, J. Löchner, and D. Burschka, "Interaction-aware probabilistic behavior prediction in urban environments," Apr. 2018, *arXiv:1804.10467*. [Online]. Available: https://arxiv.org/abs/1804.10467

[56] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," May 2018, *arXiv:1805.11773*. [Online]. Available: https://arxiv.org/abs/1805.11773

[57] E. Ohn-Bar and M. M. Trivedi, "Looking at humans in the age of self-driving and highly automated vehicles," *IEEE Trans. Intell. Veh.*, vol. 1, no. 1, pp. 90–104, Mar. 2016.

[58] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, Jun. 2011.

[59] S. Kaplan, M. A. Güvensan, A. G. Yavuz, and Y. Karalurt, "Driver behavior analysis for safe driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3017–3032, Dec. 2015.

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

[60] H.-B. Kang, "Various approaches for driver and driving behavior monitoring: A review," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, Jun. 2013, pp. 616–623.

[61] M. Kuehn, T. Hummel, and A. Lang, "Cyclist-car accidents–their consequences for cyclists and typical accident scenarios," in *Proc. 24th Int. Conf. Enhanced Saf. Vehicles*, 2015, pp. 1–8.

[62] M. Zhang, Y. Zhang, L. Zhang, C. Liu, and S. Khurshid, "DeepRoad: GAN-based metamorphic autonomous driving system testing," Feb. 2018, *arXiv:1802.02295*. [Online]. Available: https://arxiv.org/abs/1802.02295

[63] L. Neumann *et al.*, "*Nightowls*: A pedestrians at night dataset," in *Proc. Asian Conf. Comput. Vis.*, 2018, pp. 691–705. [Online]. Available: http://www.nightowls-dataset.org

[64] G. R. de Campos, P. Falcone, R. Hult, H. Wymeersch, and J. Sjöberg, "Traffic coordination at road intersections: Autonomous decision-making algorithms using model-based heuristics," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 1, pp. 8–21, Jan. 2017.

[65] J. Parkin, B. Clark, W. Clayton, M. Ricci, and G. Parkhurst, "Understanding interactions between autonomous vehicles and other road users: A literature review," Univ. West England, Bristol, England, Project Rep., 2016. [Online]. Available: http://eprints.uwe.ac.uk/29153

[66] I. Kotseruba, A. Rasouli, and J. K. Tsotsos, "Joint attention in autonomous driving (JAAD)," Sep. 2016, *arXiv:1609.04741*. [Online]. Available: https://arxiv.org/abs/1609.04741

[67] A. Rasouli and J. K. Tsotsos, "Joint attention in driver-pedestrian interaction: From theory to practice," Feb. 2018, *arXiv:1802.02522*. [Online]. Available: https://arxiv.org/abs/1802.02522

[68] P. Koopman and M. Wagner, "Autonomous vehicle safety: An interdisciplinary challenge," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 1, pp. 90–96, Jan. 2017.

[69] V. V. Dixit, S. Chand, and D. J. Nair, "Autonomous vehicles: Disengagements, accidents and reaction times," *PLoS One*, vol. 11, no. 12, 2016, Art. no. e0168054.

[70] D. Sadigh, S. Sastry, S. A. Seshia, and A. D. Dragan, "Planning for autonomous cars that leverage effects on human actions," *Robot., Sci. Syst.*, vol. 2, pp. 1–9, Jun. 2016.

[71] S. Ulbrich *et al.*, "Towards a functional system architecture for automated vehicles," Mar. 2017, *arXiv:1703.08557*. [Online]. Available: https://arxiv.org/abs/1703.08557

[72] S.-C. Lin *et al.*, "The architectural implications of autonomous driving: Constraints and acceleration," in *Proc. ACM 23rd Int. Conf. Architectural Support Program. Lang. Operating Syst.*, 2018, pp. 751–766.

[73] Mobileye. (2018). *Enabling Autonomous*. [Online]. Available: http://www.mobileye.com/futureof-mobility/mobileye-enabling-autonomous/

[74] B. Apollo. (2018). *Open Platform Details* [Online]. Available: http://apollo.auto/developer.html/

[75] S. Lefevre, D. Vasquez, and C. Laugier, "A survey on motion prediction and risk assessment for intelligent vehicles," *Robomech J.*, vol. 1, p. 1, Jul. 2014.

[76] J. Eggert, "Predictive risk estimation for intelligent adas functions," in *Proc. IEEE 17th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 711–718.

[77] C. Laugier *et al.*, "Probabilistic analysis of dynamic scenes and collision risks assessment to improve driving safety," *IEEE Intell. Transp. Syst. Mag.*, vol. 3, no. 4, pp. 4–19, Oct. 2011.

[78] M. Strickland, G. Fainekos, and H. B. Amor, "Deep predictive models for collision risk assessment in autonomous driving," Nov. 2017, *arXiv:1711.10453*. [Online]. Available: https://arxiv.org/abs/1711.10453

[79] W. Wachenfeld, P. Junietz, R. Wenzel, and H. Winner, "The worst-time-to-collision metric for situation identification," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2016, pp. 729–734.

[80] M. Elbanhawi, M. Simic, and R. Jazar, "In the passenger seat: Investigating ride comfort measures in autonomous cars," *IEEE Intell. Transp. Syst. Mag.*, vol. 7, no. 3, pp. 4–17, Jul. 2015.

[81] J. Koo, J. Kwac, W. Ju, M. Steinert, L. Leifer, and C. Nass, "Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance," *Int. J. Interact. Des. Manuf.*, vol. 9, no. 4, pp. 269–275, 2015.

[82] P. Koschorrek, T. Piccini, P. Oberg, M. Felsberg, L. Nielsen, and R. Mester, "A multi-sensor traffic scene dataset with omnidirectional video," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, Jun. 2013, pp. 727–734. [Online]. Available: http://www.cvl.isy.liu.se/research/datasets/amuse/

[83] (2018). *Apollo Data Open Platform*. [Online]. Available: http://data.apollo.auto/

[84] X. Huang, P. Wang, X. Cheng, D. Zhou, Q. Geng, and R. Yang, "The apolloscape open dataset for autonomous driving and its application," Mar. 2018, *arXiv:1803.06184*. [Online]. Available: https://arxiv.org/abs/1803.06184

[85] M.-F. Chang *et al.*, "Argoverse: 3D tracking and forecasting with rich maps," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8748–8757. [Online]. Available: https://www.argoverse.org/data.html/

[86] R. Timofte, K. Zimmermann, and L. Van Gool, "Multi-view traffic sign detection, recognition, and 3D localisation," *Mach. Vis. Appl.*, vol. 25, no. 3, pp. 633–647, Apr. 2014. [Online]. Available: http://www.vision.ee.ethz.ch/~timofter/traffic_signs/

[87] F. Yu *et al.*, "BDD100K: A diverse driving video database with scalable annotation tooling," May 2018, *arXiv:1805.04687*. [Online]. Available: https://arxiv.org/abs/1805.04687

[88] K. Behrendt, L. Novak, and R. Botros, "A deep learning approach to traffic lights: Detection, tracking, and classification," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May/Jun. 2017, pp. 1370–1377. [Online]. Available: https://hci.iwr.uni-heidelberg.de/node/6132

[89] A. Jain, H. S. Koppula, B. Raghavan, S. Soh, and A. Saxena, "Car that knows before you do: Anticipating maneuvers via learning temporal driving models," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 3182–3190. [Online]. Available: http://brain4cars.com/

[90] P. Dollár, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: A benchmark," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 304–311. [Online]. Available: http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/

[91] G. J. Brostow, J. Shotton, J. Fauqueur, and R. Cipolla, "Segmentation and recognition using structure from motion point clouds," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2008, pp. 44–57. [Online]. Available: http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamVid/

[92] R. Guzmán, J.-B. Hayet, and R. Klette, "Towards ubiquitous autonomous driving: The CCSAD dataset," in *Proc. Int. Conf. Comput. Anal. Images Patterns*. Cham, Switzerland: Springer, 2015, pp. 582–593. [Online]. Available: http://aplicaciones.cimat.mx/Personal/jbhayet/ccsad-dataset

[93] M. Cordts *et al.*, "The cityscapes dataset for semantic urban scene understanding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 3213–3223. [Online]. Available: https://www.cityscapes-dataset.com/

[94] H. Badino, D. Huber, and T. Kanade, "Real-time topometric localization," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2012, pp. 1635–1642. [Online]. Available: http://3dvis.ri.cmu.edu/data-sets/localization/

[95] E. Santana and G. Hotz, "Learning a driving simulator," Aug. 2016, *arXiv:1608.01230*. [Online]. Available: https://arxiv.org/abs/1608.01230

[96] H. Schafer, E. Santana, A. Haden, and R. Biasini. (2018). *A Commute in Data: The Comma2k19 Dataset*. [Online]. Available: https://github.com/commaai/comma2k19

[97] X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, "Spatial as deep: Spatial CNN for traffic scene understanding," in *Proc. AAAI Conf. Artif. Intell.*, 2018, pp. 1–8. [Online]. Available: https://xingangpan.github.io/projects/CULane.html

[98] M. Enzweiler and D. M. Gavrila, "Monocular pedestrian detection: Survey and experiments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2179–2195, Dec. 2008. [Online]. Available: http://www.gavrila.net/Datasets/Daimler_Pedestrian_Benchmark_D/daimler_pedestrian_benchmark_.html

[99] J. Binas, D. Neil, S.-C. Liu, and T. Delbruck, "DDD17: End-to-end davis driving dataset," Nov. 2017, *arXiv:1711.01458*. [Online]. Available: https://arxiv.org/abs/1711.01458

[100] Y. Chen *et al.*, "Lidar-video driving dataset: Learning driving policies effectively," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 5870–5878. [Online]. Available: http://www.dbehavior.net/

[101] N. Pugeault and R. Bowden, "How much of driving is preattentive?" *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5424–5438, Dec. 2015. [Online]. Available: http://cvssp.org/data/diplecs/

[102] S. Alletto, A. Palazzi, F. Solera, S. Calderara, and R. Cucchiara, "DR(eye)VE: A dataset for attention-based tasks with applications to autonomous and assisted driving," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2016, pp. 54–60. [Online]. Available: http://imagelab.ing.unimore.it/dreyeve

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

[103] (2010). *EISATS*. [Online]. Available: https://ccv.wordpress.fos.auckland.ac.nz/eisats/

[104] (2016). *Elektra Datasets*. [Online]. Available: http://adas.cvc.uab.es/elektra/datasets/

[105] A. Ess, B. Leibe, K. Schindler, and L. van Gool, "A mobile vision system for robust multi-person tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8. [Online]. Available: https://data.vision.ee.ethz.ch/cvl/aess/dataset/

[106] *FREE FLIR Thermal Dataset for Algorithm Training*. Accessed: Apr. 4, 2019. [Online]. Available: https://www.flir.com/oem/adas/adas-dataset-form/

[107] G. Pandey, J. R. McBride, and R. M. Eustice, "Ford campus vision and lidar data set," *Int. J. Robot. Res.*, vol. 30, no. 13, pp. 1543–1552, 2011. [Online]. Available: http://robots.engin.umich.edu/SoftwareData/Ford

[108] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The german traffic sign detection benchmark," in *Proc. Int. Joint Conf. Neural Netw.*, Aug. 2013, pp. 1–8. [Online]. Available: http://benchmark.ini.rub.de/

[109] S. Meister, B. Jähne, and D. Kondermann, "Outdoor stereo camera system for the generation of real-world benchmark data sets," *Opt. Eng.*, vol. 51, no. 2, 2012, Art. no. 021107. [Online]. Available: https://hci.iwr.uni-heidelberg.de/benchmarks/Challenging_Data_for_Stereo_and_Optical_Flow

[110] (2018). *HD1K Benchmark Suite*. [Online]. Available: http://hci-benchmark.org/

[111] G. Varma, A. Subramanian, A. Namboodiri, M. Chandraker, and C. V. Jawahar, "IDD: A dataset for exploring problems of autonomous navigation in unconstrained environments," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2019, pp. 1743–1751. [Online]. Available: http://idd.insaan.iiit.ac.in/

[112] Y. Choi *et al.*, "KAIST multi-spectral day/night data set for autonomous and assisted driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 934–948, Mar. 2018. [Online]. Available: https://sites.google.com/view/multispectral/

[113] J. Jeong, Y. Cho, Y.-S. Shin, H. Roh, and A. Kim, "Complex urban LiDAR data set," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2018, pp. 6344–6351. [Online]. Available: http://irap.kaist.ac.kr/dataset/

[114] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231–1237, 2013. [Online]. Available: http://www.cvlibs.net/datasets/kitti/

[115] A. Mogelmose, M. M. Trivedi, and T. B. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1484–1497, Dec. 2012. [Online]. Available: http://cvrr.ucsd.edu/LISA/lisa-traffic-sign-dataset.html

[116] J.-L. Blanco-Claraco, F.-Á. Moreno-Dueñas, and J. González-Jiménez, "The Málaga urban dataset: High-rate stereo and LiDAR in a realistic urban scenario," *Int. J. Robot. Res.*, vol. 33, no. 2, pp. 207–214, 2014. [Online]. Available: https://www.mrpt.org/MalagaUrbanDataset

[117] G. Neuhold, T. Ollmann, S. R. Bulo, and P. Kontschieder, "The mapillary vistas dataset for semantic understanding of street scenes," in *Proc. ICCV*, Oct. 2017, pp. 5000–5009. [Online]. Available: https://www.mapillary.com/dataset/vistas

[118] (2017). *The Nexar Challenge and Nexet Dataset*. [Online]. Available: https://www.getnexar.com/challenge-2/

[119] H. Caesar *et al.*, "nuScenes: A multimodal dataset for autonomous driving," 2019, *arXiv:1903.11027*. [Online]. Available: https://www.nuscenes.org/overview

[120] W. Maddern, G. Pascoe, C. Linegar, and P. Newman, "1 year, 1000 km: The oxford robotcar dataset," *IJ Robot. Res.*, vol. 36, n0. 1, pp. 3–15, 2016. [Online]. Available: http://robotcar-dataset.robots.ox.ac.uk/

[121] A. Teichman, J. Levinson, and S. Thrun, "Towards 3D object recognition via classification of arbitrary object tracks," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2011, pp. 4034–4041. [Online]. Available: https://cs.stanford.edu/people/teichman/stc/

[122] D. Pfeiffer, S. Gehrig, and N. Schneider, "Exploiting the power of stereo confidences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2013, pp. 297–304. [Online]. Available: http://www.6d-vision.com/ground-truth-stixel-dataset

[123] P. Mirowski *et al.*, "The streetlearn environment and dataset," Mar. 2019, *arXiv:1903.01292*. [Online]. Available: https://arxiv.org/abs/1903.01292

[124] M. De Deuge, A. Quadros, C. Hung, and B. Douillard, "Unsupervised feature learning for classification of outdoor 3d scans," in *Proc. Australas. Conf. Robot. Autom.*, vol. 2, 2013, pp. 1–9. [Online]. Available: http://www.acfr.usyd.edu.au/papers/SydneyUrbanObjectsDataset.shtml

[125] C. Caraffi, T. Vojíř, J. Trefný, J. Šochman, and J. Matas, "A system for real-time detection and tracking of vehicles from a single car-mounted camera," in *Proc. 15th Int. IEEE Conf. Intell. Transp. Syst.*, Sep. 2012, pp. 975–982. [Online]. Available: http://cmp.felk.cvut.cz/data/motorway/

[126] C. Wojek, S. Walk, and B. Schiele, "Multi-cue onboard pedestrian detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 794–801. [Online]. Available: https://tinyurl.com/ycyp62bq

[127] (2018). *Tusimple Benchmark*. [Online]. Available: http://benchmark.tusimple.ai/#/

[128] E. Romera, L. M. Bergasa, and R. Arroyo, "Need data for driver behaviour analysis? Presenting the public UAH-DriveSet," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 387–392. [Online]. Available: https://tinyurl.com/yd8f5maz

[129] (2016). *Udacity Self-Driving Car*. [Online]. Available: https://github.com/udacity/self-driving-car

[130] *UTBM Multisensor Ros-Based Dataset for Autonomous Driving*. Accessed: Apr. 4, 2019. [Online]. Available: https://epan-utbm.github.io/utbm_robocar_dataset/

[131] H. Yin and C. Berger, "When to use what data set for your self-driving car algorithm: An overview of publicly available driving datasets," in *Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2017, pp. 1–8.

[132] S. Wang *et al.*, "Torontocity: Seeing the world with a million eyes," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 3028–3036.

[133] D. Ramachandram and G. W. Taylor, "Deep multimodal learning: A survey on recent advances and trends," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 96–108, Nov. 2017.

[134] M. Bojarski *et al.*, "End to end learning for self-driving cars," Apr. 2016, *arXiv:1604.07316*. [Online]. Available: https://arxiv.org/abs/1604.07316

[135] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding data augmentation for classification: When to warp?" Sep. 2016, *arXiv:1609.08764*. [Online]. Available: https://arxiv.org/abs/1609.08764

[136] A. Gaidon, Q. Wang, Y. Cabon, and E. Vig, "Virtual worlds as proxy for multi-object tracking analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 4340–4349.

[137] G. Ros, L. Sellart, J. Materzynska, D. Vazquez, and A. M. Lopez, "The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 3234–3243.

[138] S. R. Richter, Z. Hayder, and V. Koltun, "Playing for benchmarks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 2, Oct. 2017, pp. 2213–2222.

[139] I. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[140] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from simulated and unsupervised images through adversarial training," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, vol. 2, no. 4, pp. 2107–2116.

[141] H.-P. Schöner, "Simulation in development and testing of autonomous vehicles," in *Proc. Internationales Stuttgarter Symp.* Wiesbaden, Germany: Springer, 2018, pp. 1083–1095.

[142] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proc. Conf. Robot Learn.*, 2017, pp. 1–16.

[143] S. Shah, D. Dey, C. Lovett, and A. Kapoor, "AirSim: High-fidelity visual and physical simulation for autonomous vehicles," in *Field and Service Robotics*. Cham, Switzerland: Springer, 2018, pp. 621–635.

[144] *Apollo Simulation*. Accessed: Oct. 15, 2018. [Online]. Available: http://apollo.auto/platform/simulation.html

[145] *Nvidia Drive Constellation-Virtual Reality Autonomous Vehicle Simulator*. Accessed: Oct. 23, 2018. [Online]. Available: https://www.nvidia.com/en-us/self-driving-cars/drive-constellation/

[146] *Pegasus Research Project*. Accessed: Nov. 15, 2018. [Online]. Available: https://www.pegasusprojekt.de/en/home

[147] A. Pütz, A. Zlocki, J. Bock, and L. Eckstein, "System validation of highly automated vehicles with a database of relevant traffic scenarios," in *Proc. 12th ITS Eur. Congr.*, vol. 1, 2017, p. E5.

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.

[148] S. Ebrahimi, A. Rohrbach, and T. Darrell, "Gradient-free policy architecture search and adaptation," Oct. 2017, *arXiv:1710.05958*. [Online]. Available: https://arxiv.org/abs/1710.05958

[149] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[150] N. Japuria, G. Habibi, and J. P. How, "CASNSC: A context-based approach for accurate pedestrian motion prediction at intersections," in *Proc. 31st Conf. Neural Inf. Process. Syst. (NIPS)*, 2017, pp. 1–8.

Junyao Guo received the bachelor's degree in electronic engineering from Tsinghua University in 2013 and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University in 2018. She is currently a Senior Research Scientist with the North American Research and Development Center, Advanced AI Group, LG Silicon Valley Lab. Her research interests include optimization and distributed algorithms for machine learning, with applications in smart grids and the IoT.

Unmesh Kurup received the bachelor's degree in computer science and engineering from the Cochin University of Science and Technology and the Ph.D. degree in computer science and engineering specializing in artificial intelligence from The Ohio State University. He was a Senior Scientist with the Bosch Center for Artificial Intelligence, where he led data science projects in variety of domains, including energy, healthcare, and ADAS. He is currently a Senior Manager with the North American Research and Development Center, Advanced AI Group, LG Silicon Valley Lab, where he is involved in inference and learning at the edge. He has authored the theory of expectation-driven cognition (EDC), which models human intelligence as a duality between perception and prediction. His research interests include the intersection of machine learning, big data, and human cognition.

Mohak Shah is currently a Technology Leader and a Technologist with an established history of leading high-impact initiatives in AI, the IoT, digital transformation, and building data-driven businesses and strategies. He is the VP of AI and Machine Learning at LG Silicon Valley Lab. He also holds an adjunct faculty position with the University of Illinois at Chicago. His experience in leading technology initiatives spans a broad range of verticals, including autonomous driving, automotive, aviation, energy, and healthcare leading diverse teams from research, software, and businesses. As a Machine Learning and AI Scientist, he has developed novel technologies with high-impact business applications. He has authored *Evaluating Learning Algorithms: A Classification Perspective* (Cambridge University Press) and has published more than 45 research articles in top conferences and journals in the analytics space and patented technologies. He regularly contributes to the community through participation in scientific panels, organizing committees, review committees, and social impact programs. He was the General Chair of the ACM SIGKDD 2016 Conference.

Authorized licensed use limited to: Queen's University. Downloaded on February 21,2022 at 05:29:28 UTC from IEEE Xplore. Restrictions apply.