

Received December 31, 2021, accepted February 7, 2022, date of publication February 24, 2022, date of current version March 9, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3154419

CaT: CAVS Traversability Dataset for Off-Road Autonomous Driving

SUVASH SHARMA^{1,2}, (Member, IEEE), LALITHA DABBIRU¹, TYLER HANNIS¹,
GEORGE MASON¹, DANIEL W. CARRUTH¹, MATTHEW DOUDE³, CHRIS GOODIN¹,
CHRISTOPHER HUDSON¹, SAM OZIER², JOHN E. BALL², (Senior Member, IEEE),
AND BO TANG², (Senior Member, IEEE)

¹Center for Advanced Vehicular Systems, Mississippi State University, Mississippi State, MS 39762, USA

²Electrical and Computer Engineering Department, Mississippi State University, Mississippi State, MS 39762, USA

³Dynetics Inc., Huntsville, AL 35806, USA

Corresponding author: Bo Tang (tang@ece.msstate.edu)

ABSTRACT In the context of autonomous driving, the existing semantic segmentation concept strongly supports on-road driving where hard inter-class boundaries are enforced and objects can be categorized based on their visible structures with high confidence. Due to the well-structured nature of typical on-road scenes, current road extraction processes are largely successful and most types of vehicles are able to traverse through the area that is detected as road. However, the off-road driving domain has many additional uncertainties such as uneven terrain structure, positive and negative obstacles, ditches, quagmires, hidden objects, etc. making it very unstructured. Traversing through such unstructured area is constrained by a vehicle's type and its capability. Therefore, an alternative approach to segmentation of the off-road driving trail is required that supports consideration of the vehicle type in a way that is not considered in state-of-the-art on-road segmentation approaches. To overcome this limitation and facilitate the path extraction in the off-road driving domain, we propose traversability concept and corresponding dataset which is based on the notion that the driving trails should be finely resolved into different sub-trails and areas corresponding to the capability of different vehicle classes in order to achieve safe traversal. Based on this, we consider three different classes of vehicles (sedan, pickup, and off-road) and label the images corresponding to the traversing capability of those vehicles. So the proposed dataset facilitates the segmentation of off-road driving trail into three regions based on the nature of the driving area and vehicle capability. We call this dataset as CaT (CAVS Traversability, where CAVS stands for Center for Advanced Vehicular Systems) dataset and is publicly available at <https://www.cavs.msstate.edu/resources/downloads/CaT/CaT.tar.gz>.

INDEX TERMS Traversability, autonomous driving, semantic segmentation, off-road, dataset.

I. INTRODUCTION

Either from the point of research or from the point of its applicability in fields such as robotics, surveillance, military, etc., autonomous driving in off-road environment is gaining increased attention nowadays. In off-road environments, semantic segmentation [1]–[4] is often used for understanding of scenes around the vehicle. Per-pixel segmentation assigns labels to each pixel in a frame of data according to detection and classification of objects in the scene. In the case of on-road driving, there are clear boundaries for the road. The segmented road is equally drivable by all on-road

vehicles irrespective of their specific types and capabilities. However, for off-road driving, the off-road trail presents many challenges such as obstacles, ditches, quagmires, logs across the road, and more. A vehicle's ability to negotiate these challenges is dependent on numerous factors of the vehicle's design. The traversability of a trail is still uncertain without the consideration of the vehicle's capabilities. Thus, the driving decision - whether to drive through the specific area or not - should be made on the basis of the maneuvering vehicle type and its capability.

With the advent of deep learning (DL), highly accurate output can be easily achieved from scene understanding algorithms. Such algorithms are data-driven and typically need very large training datasets. To fulfill the data

The associate editor coordinating the review of this manuscript and approving it for publication was Eduardo Rosa-Molinar¹.

requirements, several datasets for autonomous driving have been proposed. However, most publicly available datasets are targeted for on-road vehicles in well-structured environments [5]–[10]. There are very few datasets for unstructured environment [11], [12] despite potential applications such as rural driving [13], robot-based forestry studies [14], surveillance [15], etc. The available datasets have been labeled using standard pixel segmentation that identify classes of objects in a scene but do not consider the relationship between types of challenges on off-road trails and the capabilities of different vehicles. To address these limitations, we propose a segmentation approach, and corresponding dataset, that is based on the concept of traversability that is mediated by class of vehicle.

In the context of traversability assessment of the off-road or other unstructured area, several works can be noted. Gresenz et.al in [16] provide roughness-based road area classification dataset generated from GPS (Global Positioning System), IMUs (Inertial Measurement Units), and wheel rotation sensor. However, this lacks the image-based human judgmented dataset. A traversability analysis in aerial imageries with height assisted dataset has been proposed in [17] targeted for three different types of robots (wheeled, tracked, and legged). However, we believe it would give a very rough estimation of traversability only from height since there could be some objects that can be easily override even though their height is above some threshold. Camera images-based traversability assessment of the construction area is proposed in [18]. However, this dataset considers the only one type of excavator and is not applicable for driving scenes. Different from these contributions, our CaT dataset is based on expert-level human driver's visual judgement of the trails. Furthermore, it correlates the diverse trail area to the capability of the vehicle that is quantified and verified by their overriding force.

A. TRAVERSABILITY AND ITS NECESSITY

In this work, we refer the drivability of a given off-road trail based-on the vehicle capability as traversability. So, the decision regarding whether a vehicle-type can traverse through an off-road trail with minimum or no-damage is defined under traversability concept [19]. Further, if the given vehicle can pass the trail with minimum or no-damage, we call the trail as a traversable, and if not, it is non-traversable for that vehicle. The area that is non-traversable for one vehicle-type could be traversable for the other. So, traversability solely cannot be defined as a function of the nature of driving-trail. It also depends on the type of corresponding maneuvering vehicle as different vehicle-class have different traversing capability. Be it for on-road or off-road environment, the existing datasets consider that the parts of the segmented scene assigned as the road/track are equally drivable irrespective of the road nature and vehicle type. This consideration holds strongly in case of the on-road autonomous vehicles. However, for the off-road case, the road (or driving trail) is not uniform and is full of irregularities due to uneven terrain, ditches, steep slopes, logs, bushes, shrubs etc.



FIGURE 1. Example of traversability assessment in camera images. Left: original image, right: traversability labels overlaid onto the image where red, green and blue colors correspond the traversable areas by off-road, pickup, and sedan type vehicles. In this specific example, due to the absence of other irregularities on the road, the only difference between the road parts that could be traversed by each vehicle type is in-terms of the extent on two sides of the road it can travel.

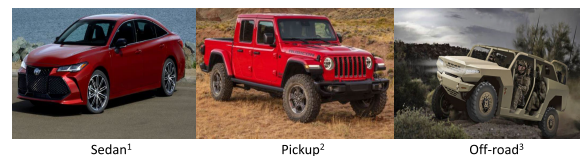


FIGURE 2. Vehicle-types considered for the traversability labels.

Due to such irregularities, the capability of the running vehicle to traverse-through is skeptical. Such capability to traverse through the adverse scenarios on the off-road environment is subjective to the type of the vehicle. For example, A pickup could traverse through the road overriding a medium-sized log (considered as an obstacle) that cannot be override by a sedan. Similar scenario may arise in case of shrubs and bushes that usually invade the part of driving trail from both sides. So, in this paper we propose that for the road segmentation in off-road driving environment, the extraction of road tracks that are based on the existing segmentation datasets do not correlate the nature of the unstructured driving. However, the fine partitioning of the road-based on the traversing capability of several vehicles would be helpful for the successful driving. To demonstrate and strengthen the concept, this paper includes several experiments and assessments to define whether the camera imagery could be used to define the traversability for three different types of vehicles.

To demonstrate the above concept and help the research community, we propose a new dataset for semantic road segmentation with fine partitioning of off-road autonomous driving trail based on the traversing capability of three different vehicles. We call it the CaT dataset. From the road segmentation point of view, the proposed dataset is comparable to Kitti [5] dataset where only the road in the image is labelled leaving other parts of the scenes -for example, forest, vegetation, sky, etc.- unlabelled. However, the road extraction problem is not formulated on the basis of binary classification as the road-parts themselves are to be classified into various sub-classes to correspond the traversability concept. Such sub-classes are dependent on vehicle type (e.g. sedan, pickup or off-road) traversing along that particular part of the off-road area. While doing so, we consider the vehicular strengths and their traversing nature in-terms of overriding any unintended obstacles, shrubs, bushes etc.

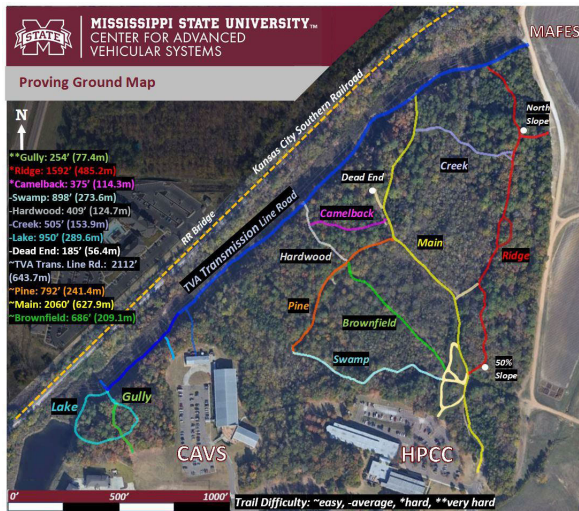


FIGURE 3. Data collection map. Note that the acronyms representations as: TVA: Tennessee Valley Authority. Best viewed in color.

Also it should be ensured that the vehicle is undamaged after traversing through that area. As shown in Figure 1, the edge of red color represents the extent which off-road vehicle could traverse through, green color represents that for a pickup, and the blue color represents that for a sedan at both ends. The corresponding vehicle-types (Sedan,¹ Pickup,² Humvee³) for these levels are shown in Figure 2.

In Figure 3, we show our data collection sites located in front of HPCC (High Performance Computing Colaboratory, a Mississippi State University's research extension) building. We basically collect our data from three trails namely the main trail, the powerline trail, and the brownfield trail that are highlighted by yellow, blue, and green colors respectively. These three trails have the approximate lengths of 0.4 miles, 0.51 miles, and 0.13 miles, respectively. The images are collected with the consideration of dark and enhanced lighting condition with different camera filters possible with the Sekonix SF3325-100 camera model [20]. All the images are labeled as per the traversability concept with following assumptions:

- All the assigned labellers can accurately delineate the driving reasons based on the difficulty level corresponding to the vehicle capability of driving as per the rule: Off-road > Pickup > Sedan.
- The traversability labels are to be assigned solely based on the camera images. To accurately delineate the regions, images should be zoomed such that it helps to precisely notice the possible uncertainties and tell

¹Image accessed from <https://www.caranddriver.com/features/g22344863/full-size-sedans/?slide=7>

²Image accessed from <https://cars.usnews.com/cars-trucks/best-pickup-trucks>

³Image accessed from <https://www.autoweek.com/news/future-cars/a34549967/kia-will-beef-up-its-military-lineup-with-humvee-style-models/>

whether they are traversable or not by the corresponding vehicle class.

- It is considered that the vehicle corresponding to the traversability label can pass the corresponding region with negligible or no damage.

The major contributions of this paper are as follows:

- A new traversability concept regarding segmentation of off-road trails is proposed in order to assess the drivability of the area using visual imageries.
- A new type of dataset, called the CaT dataset, is proposed based on the concept that whether the driving trail is actually drivable for the corresponding vehicle. We consider different vehicle classes and label the driving trail based on which part is drivable by which vehicle-type.
- We verify our labelled dataset with real-field experiment assisted by some calculations using the vehicle specifications regarding their overriding strengths, and the amount of force that a vegetation/tree can exert on the vehicle. That means, we make sure that the force exerted by the vegetation is lower compared to the overriding strength of the corresponding vehicle in order to be traversable.
- As the several environmental factors and locations affect the performance of segmentation algorithms, we consider this fact strongly while collecting images. We consider two different weathers (winter of 2018 and summer of 2019), different time of the day and different geographical location such that our dataset captures as large variations as possible.

II. RELATED WORK

A. SEMANTIC SEGMENTATION IN ON-ROAD/OFF-ROAD ENVIRONMENT

With 20 different classes for segmentation corresponding to indoor and outdoor objects like birds, tables, airplanes etc., the pascal-VOC [6] dataset sets a good practical benchmark. For the autonomous driving application, [5] provides the real-world dataset for road detection in structured road environment. In [7] and [8], city-driving-environment based complete datasets are proposed with classes like buidings, roads, vehicles, poles, etc.

Even though there are several popular datasets available for semantic segmentation based on-road driving, very few dataset are available for unstructured off-road case. In [11], the off-road semantic segmentation dataset are provided that are collected at 20Hz with 1024×768 pixel images on three different days and variable lighting conditions. The images provided are in the different formats such as RGB, NIR, and depth-images. Six different classes are considered: Obstacle, Trail/road, Sky, Grass, Tree, and Vegetation. A multimodal dataset with LiDAR and camera images annotations for segmentation in off-road domain is provided in [12] that presents challenges to existing algorithms for class imbalance and environmental topography. In terms of targeted problem and

nature of the domain, our data overlaps to some extent with both of these datasets. However, as our proposed dataset is the first of its' type, it has several differences conceptually and technically. First, this dataset is based on the assumption that the off-road semantic segmentation should be modeled as the function of the driving trail's nature and the capability of the maneuvering vehicle. Second, we focus only on the road for segmentation and ignore the background scenes and objects in the image. Furthermore, the road itself is partitioned into three different levels based on the extent to which each vehicle could traverse.

B. VEGETATION OVERRIDING

Studies regarding vehicle maneuverability in non-ideal terrain have been performed for decades [21], [22] and have matured to provide a comprehensive understanding of vehicle physics a variety of areas and terrains [23]. Despite modern navigation algorithm's sophistication, it might not always be possible to plot a route that avoid collisions with tall bush and small trees. However, from the vegetation override formula and terrain surveys of bio-types and densities, [24] states that it can be possible to calculate the force override needed for a ground vehicle to traverse through. A vehicle's capability to traverse through an area is not the sole concern when optimizing a vehicle for its terrain. Terrain analysis can also allow for a control system to account for the vehicle's velocity traveling through areas providing an addition optimization when route planning [25].

III. DATA COLLECTION AND LABELLING PLATFORMS

We provide the brief overview of collection methodologies, the hardware setup for data collection along with the introduction with the labeling platform. As our proposed traversability concept is based on the camera images only, we limit the description of camera only, even though we have a full sensor setup.

A. VEHICLE

The vehicle we used in our data collection platform along with the sensors placed over it is as shown in Figure 4. We chose the Polaris Ranger crew XP100 due to its' good ground clearance and stronger chassis with full-body skid plate that is quite favourable for off-road driving. All the sensors are placed on a special wooden structure which we call 'Ranger hat'. In Ranger hat, the lidar and Global Navigation Satellite System/Inertial Navigation System (GNSS/INS) modules are placed vertically up and down. Two cameras are placed on both the sides of GNS/INS module symmetrically. A zoomed-in view of the Ranger hat is as shown in Figure 4.

B. CAMERA

As shown in Figure 4 (b), two white modules on both the sides of yellow IMU unit on the ranger hat are the camera sensors we used. These are Sekonix SF3325-100 model with RCCB (Red-Clear-Clear-Blue) color filters and AR0231 CMOS image sensor with an active-pixel array

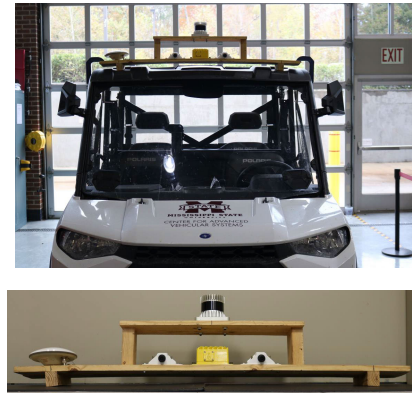


FIGURE 4. Vehicle used for data collection and sensors position. Top: Ranger with sensors placed on it. Bottom: zoomed-in view of the sensor setup.

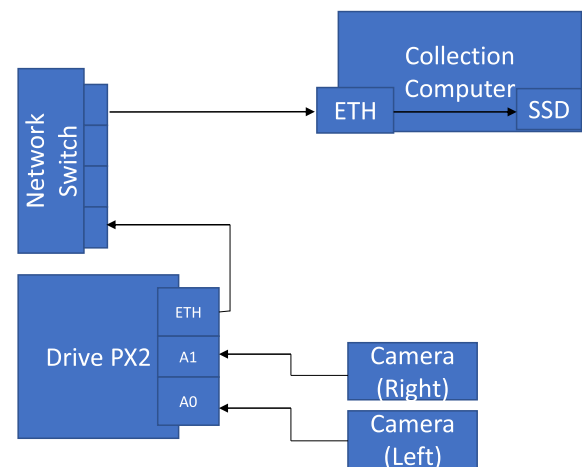


FIGURE 5. Our data collection pipeline with connection details. ETH = Ethernet. SSD = Solid State Drive. A0 and A1 are PX2 camera input ports.

of 1928×1208 with a LED Flicker Mitigation (LFM). With IP69k rating these cameras are capable to resist the effect of adverse environmental conditions like high temperature, dust, and high water pressure. The serialized input-output is supported with MAX96705 GMSL serializer [20] with 27MHz.

C. DRIVE PX 2

Nvidia's Drive PX2 also holds special position in our collection platform. As shown in Figure 5, PX2 module resides in between cameras output and network. Drive PX2 provides a powerful and easy platform for the autonomous driving-based algorithms and are mostly used in Tesla's enhanced autopilot vehicles [26].

D. COLLECTION PIPELINE

The overall collection pipeline along with the connection detailing is shown in the Figure 5. As shown in the figure, both of the cameras are connected to the ROS (Robotic Operating System) nodes available on the Drive PX2.

The output of the camera is transferred over the network through PX2 and are finally saved in the collection computer. All the dataset are collected in rosbag formats which we transfer to the CAVS servers for the final storage. Representative images from each of the trail are selected from the rosbag extractions and provided to the labelers to annotate with the traversability labels.

IV. FIELD ASSESSMENT FOR TRAVERSABILITY

In this section, we explain the experiment performed in real driving trail in order to assess the traversability before handing over the labeling task to the labelers. For this assessment, we conduct two experiments: measuring the tree density and calculating the overriding force, based on which the traversability labels are assigned as per the strength and driving capability of each type of vehicle. Overriding force refers to the amount of force required to pass through an area without avoiding the obstacles, tree-logs, trees, etc. In other words this is the force that a tree-trunk exerts on the vehicle that needs to overcome and pass-through. Note that, the overriding force cannot be defined for all type of obstacles/uncertainties. For example, even though there is no such study to define quantitatively, the overriding force for a tall and wide tree-trunks cannot be defined and hence the traversability concept is valid for only the area around it.

The movement of a vehicle down a forested trail is restricted based on the weight, width, and type of vehicle. Measurements of tree diameter and spacing along the length of a traverse were made to objectively estimate traversability of a section of trail, based on the vehicle type. With the collected measurements, we generate the ground truths for certain regions with vegetation overriding force equations. After generalizing the outcomes of the experiment performed in this section, the labelers were instructed to assign the traversability labels for each vehicle type with high confidence.

A. MEASURING TREE DENSITY

A survey of three trails on the CAVS proving grounds was performed at several locations selected to represent the vegetation for that area of the trail. These surveys collected data on the diameter of any vegetation at 36" from the ground (approximately waist level) via calipers. Note that the vegetations with height more than this reach beyond the vehicle's height and are exempted from traversability definition. The surveys took place in a 12'x12' grid on both the left and right sides of the data collection vehicle since the vegetation beyond this was dense and mostly exceeding the height of 36". The thickness of any vegetation at least one meter tall was recorded at one meter off the ground and recorded into a grid referenced away from the corners of the data collection vehicle. An example of this assessment can be visualized in second row of Figure 6. This process was repeated to collect data at 9 separate and diverse locations.

B. OVERRIDING THE VEGETATIONS AND VERIFICATION OF TRAVERSABILITY LABELS

Equation 1 gives the maximum force (F) to override a tree of diameter (d) and (h) is the height of the bumper or push-bar [24], [27]. The mass of the vehicle and velocity are compared to this force to determine the override. Note that the unit of d and h are inches and that of F is Kilo-newtons.

$$F = (10.86 - 0.0534h)d^3 \quad (1)$$

The average force to override vegetation, as defined in equation 2 [27], is a motion resistance summed with other resistance forces such as slope.

$$F = 2.62 * d^3 \quad (2)$$

Utilizing the tree diameter over the course of the entire traverse, the maximum force override was computed for each tree diameter using equation 1. In the third row of Fig. 6 the illustration of the distance from the edge of the road and the average force required for overriding is demonstrated in graphical form. Each of the three vehicle systems are presented in an effort to objectively quantify the maximum traversing capability in both the sides of the trail as a function of the vehicle-type.

V. LABELING

Four undergraduate and one graduate student were assigned for the labeling task. An exploratory tour of the collection site was taken by all the labelers. The labelers were experienced in driving sedan and pickup types and were able to judge what level of difficulty could a vehicle can override in the driving trail. Compared to sedan and pickup, the undergraduates labelers were less experienced with off-road vehicle. They were given a test drive around the data-collection site and special instructions about it and nature of the areas it can cover while driving. Furthermore, the actual concept of traversability as we define in this paper was made available before starting the labeling process.

As an annotating tool, BasicAI [28] was used as it provides easy handling for a group of labelers as per the team administrator assigning which members are responsible for which image. In this platform, the administrator can check the individual's work and give instant feedback. A single expert with the deeper understanding of all three vehicle-types and sufficient off-road driving experience was assigned to audit all the labelled images. Each image in the released dataset is checked and verified by the auditor for quality and consistency.

As we consider three different vehicles, we pose this dataset as a three-class dataset. The classes of vehicles considered are: Sedan type (front-wheel drive sedan with standard wheels and tires), Pickup type (4 × 4 pickup truck with standard wheels and tires), off-road vehicle type (side-by-side all terrain vehicle that have all-wheel drive, off-road wheels and tires). In figure 7, we show some examples of images and corresponding traversability labels. Each class label and the corresponding RGB values assigned to it is as shown in

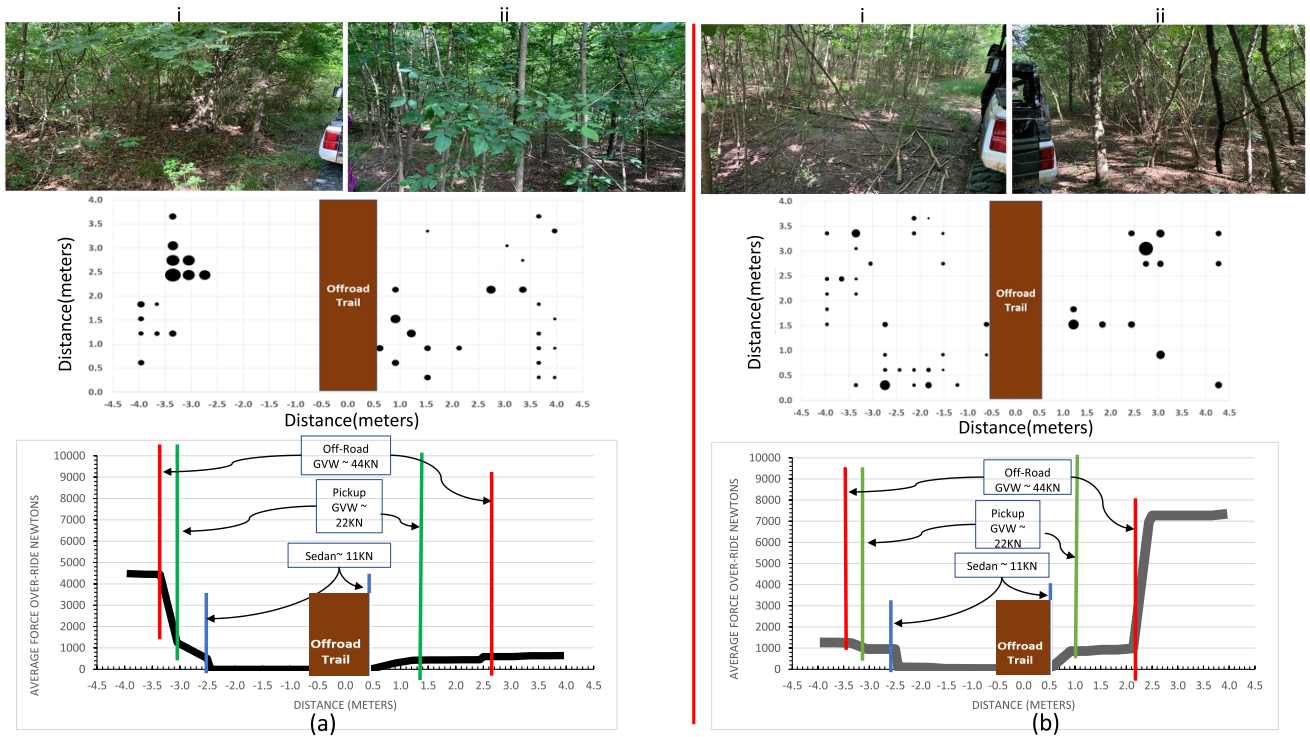


FIGURE 6. Visualization of tree density plot and corresponding force required to override the area. (a) is from main trail and (b) is from brown field. Further, note that the first row is the original images towards the left (labeled as i) and right (labelled as ii) side of the off-road trails, second row shows the graph representing the observed tree density on either side of the trail (note that the bold circles represent the measurable tree location and are not to scale), and third row is the demonstration of required amount of force needed to override the terrain with respective ground vehicle.

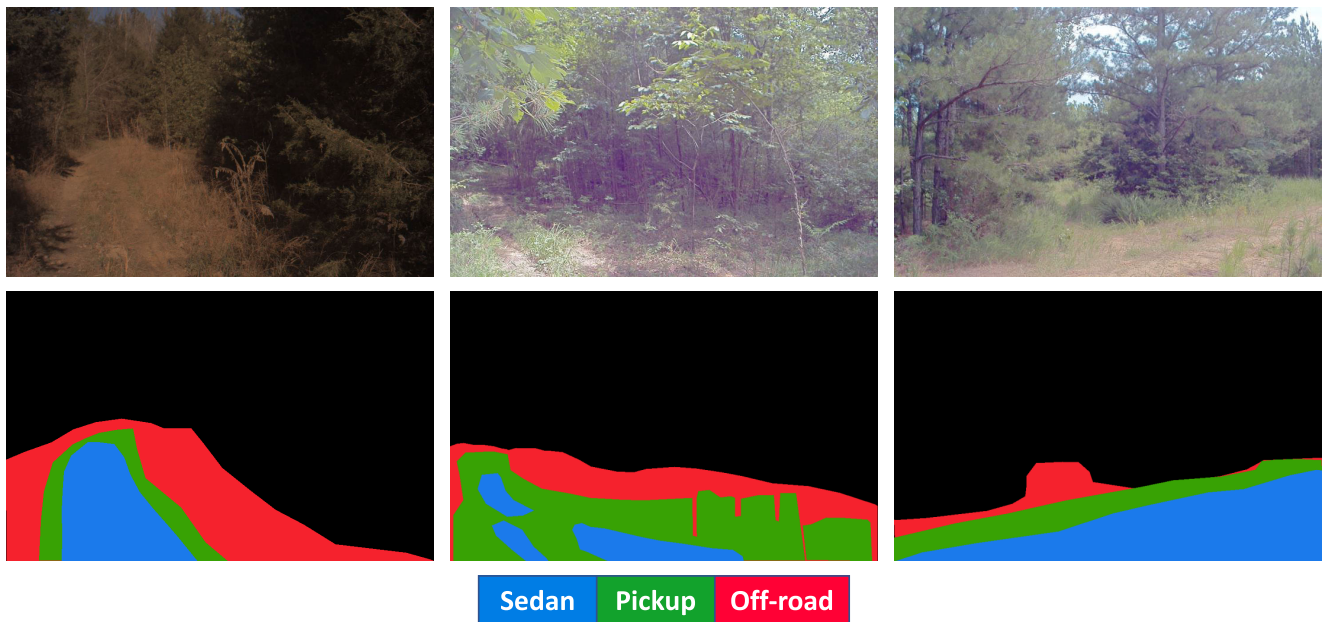


FIGURE 7. Example images and annotations from the proposed dataset. Best viewed in color. Note that each vehicle-type includes several vehicles with generalized features defined as: sedan: front-wheel drive with standard wheels and tires, pickup: 4 x 4 trucks with standard wheels and tires, off-road: side-by-side all terrain vehicle that have all-wheel drive with off-road wheels and tires.

the color chart. Each label corresponds to the area that could be covered by each vehicle type. It is worthy to note that the traversability labels are assigned based on the traversing

capability of each vehicle defined by the hierarchy: off-road > pickup > sedan. So the part of the trail labeled with lower rank in the hierarchy is also traversable by the vehicle in



FIGURE 8. Comparing traversability labels by different labelers. First column is the images given to each labeler and each columns represent the labelled images by each labelers with name on the top.

higher level. That is, the one with the blue label is traversable by all three vehicle type, the one with the green label is traversable by both pickup truck and off-road type, and the one with red label is only traversable by the off-road-type only.

On the basis of traversing area and each vehicle-type-capability, we set the following common guidelines to be followed by each labeler:

- Draw each vehicle’s complete traversable area. The labeling polygons may or may not overlap.
- Assume the sedan is capable of driving on smooth ground, gravel roads, etc.
- Assume that the off-road vehicle is capable of driving over brush up to three feet high.
- Trace around trees and other obstacles if they are not traversable.
- Labeling should be performed at approximately 150-200% zoom level so as to avoid confusion over the boundaries as much as possible.

A. HOW CLOSE ARE OUR LABELERS?

In case of off-road scenario, driving is more subjective task that depends upon the various aspects like driver’s judgement, their experience in-terms-of time and vehicle type, and their driving habits which could be categorized as a rough or smooth driver. Based on this, even though we provided the detailed labeling guidelines with site tour, the labeling task for traversability concept is influenced by the individual perception that varies among the labelers. In this section, we briefly assess and visualize how closely the environment had been judged by our labelers who have different driving experiences.

As seen in the Figure 8, our labelers have the common judgement of the driving environment and they agree upon the traversability labels when the areas are fairly far from the probable class-boundaries. However, as per the experience and personal judgement of each labeler, slight differences come around the boundaries. These differences are solely due to the experience of the labeler and their driving nature. We observed no other major differences in the labels. In order

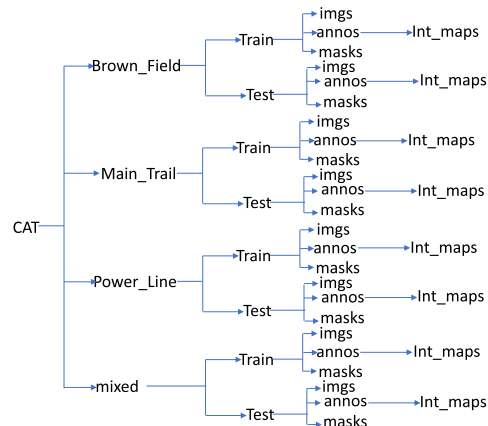


FIGURE 9. Tree structure of the data storage directory.

TABLE 1. Summary of data statistics under CaT dataset.

Dataset	no. of images	Classes	%
CaT	1812	Sedan	21.95%
		Pickup	12.26%
		off-road	20.67%
		background	45.12%

to quantify the differences among the labelers, we calculate some segmentation metrics that represent the pixel level overlaps. While doing so we consider labels from one labeler as reference or ground-truth and that of another as predicted images. The average metrics based on global accuracy, class average accuracy, and mean intersection over union [3] are 83.84%, 81.65%, and 68.90% respectively.

VI. DATASETS

A. FORMAT

The tree structure of our database is as shown in the Figure 9. The “imgs” folder contains the original images and “annos” folder contains the corresponding annotations. The “masks” folder contains the same images inside “annos” folder but the background is masked. The “int_maps” folder contains the integer label assigned images and are in png format.

B. DATA GROUPING AND STATISTICS

The pool of 1812 images is separated into training and testing set. Out of all the labelled images, 70% of the images are grouped as training set and remaining 30% are grouped as testing set. While selecting the testing set, we strictly consider the variability among the images such that the testing images also cover the general and special overview of all the trails considered. Similar separation is also provided for each of the driving trail -Brown Field, Main Trail, and Powerline-such that users could train their model in one trail and test in another trail to evaluate their model’s generalizing behaviour. Table 1 shows the pixel percentage of each class in the proposed dataset.

TABLE 2. Currently assigned RGB values to each class for each dataset items and suggested IDs for each class.

Dataset	Class	RGB value	ID
CaT	Off-road	(245,34,45)	3
	Pickup	(56,162,4)	2
	Sedan	(27,122,235)	1
	Background	(59,93,4)	0

VII. DATABASE ACCESS AND USAGE

Our complete data-pack containing all the datasets described above is available in <https://www.cavs.msstate.edu/resources/downloads/CaT/CaT.tar.gz>. We provide this in the compressed .tar file named as CAT.tar.gz and is open-sourced. As shown in Table 2, we suggest to change the RGB values of each class to corresponding grayscale values given by IDs column to perform training. However, we also provide this format inside the “int_maps” folder.

VIII. TRAINING WITH TRAVERSABILITY DATASET

In this section we explain the usage of proposed CaT dataset with the details including pre-processing, training, and testing results. Further, we analyze the areas where our model fails to assign traversability labels accurately. Also, we test our model with completely new scenarios to assess its’ generalization capability and present the visual results.

A. TRAINING DETAILS AND RESULTS

In order to demonstrate the road extraction in off-road driving domain, we train models designed for semantic segmentation. Technically from the training point of view, the proposed traversability concept and semantic segmentation align closely to each other. So we can use the same deep-learning architectures to train with traversability dataset as those used for semantic segmentation. So while training, we consider each class (sedan, pickup, and off-road) as independent class to be segmented. Each pixel is mapped into a unique integer value as shown in Table 2. So this training approach is an alias of what is done in semantic segmentation. However, considering the practical scenario, the output of the trained model should be analyzed based-on the vehicle traversing capability such that the part belonging to sedan class is also traversable by both pickup and off-road.

We train different versions of pyramid scene parsing network [30] (PSPNet) in order to monitor the performance on the proposed traversability dataset. We train this network for 80 epochs with a initial learning rate of 0.01 decreased at each epoch based on poly-learning rate schedule [31] using SGD [32]. All the related training experiments are performed in NVIDIA Quadro GP 100 GPU using PyTorch. As shown in Table 3, we use Intersection over union (IoU) as a metric to analyze the performance of the trained model in CaT test set. Considering C be the total number of classes, n_{ij} be the number of pixels from true class i that lie in predicted class j , n_i be the total number of pixels lying in true class i and n be

TABLE 3. Performance of PSPNet trained on CaT training set and tested on CaT test set. Note that the entries under sedan, pickup and off-road header represent the class-wise IoU and mIoU represents the mean of all three classes. The “Backbone” column represents the baseline network on which the PSPNet is built upon. Further, Resnet-N represents resnet [29] version with N number of layers.

Dataset	Network	Backbone	sedan	pickup	off-road	mIoU
CaT	PSPNet [30]	Resnet-18	90.44	66.62	79.71	78.92
CaT	PSPNet [30]	Resnet-34	91.21	68.64	80.52	80.12
CaT	PSPNet [30]	Resnet-50	90.70	67.40	80.00	79.36
CaT	PSPNet [30]	Resnet-101	91.64	69.08	81.00	80.57

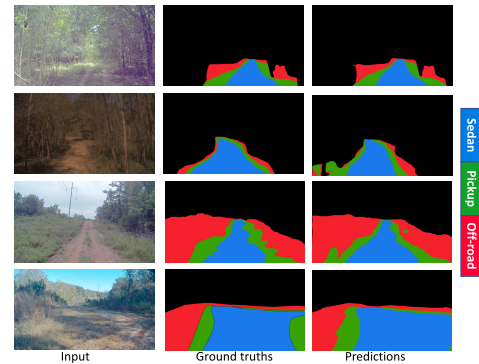


FIGURE 10. Visualizations of test operation on CaT test dataset using PSPNet-101 network. Best viewed in color.

the total number of pixels considering all the classes, the IoU and its mean of it is defined as:

- 1) IoU: It is the measurement ratio of true positives to the sum of true positives, false positive, and false negatives. When it is calculated for each class it is called Classwise IoU (CW IoU) and the mean of it is called Mean Intersection over Union (mIoU). It is expressed as:
$$IoU = \sum_i \frac{n_{ii}}{n_i + \sum_j n_{ji} - n_{ii}}$$

In Table 3, the performance of different versions of PSPNet [30] is shown. We can say that a very good value of mIoU (=80.57%) is obtained for CaT testing set with PSPNet built on top of Resnet-101 [29]. Analyzing the Class-Wise (CW) IoU, comparing the sedan, pickup, and off-road class, the recognition accuracy for pickup label is low, highest for sedan and medium for off-road. The pickup-based pixels correspond the intermediate region between smoother road and rough trails with vegetations and bushes/shrubs. Due to this, the network could incorporate these pixels towards either sedan or off-road-based regions. Furthermore, this class slightly shares the properties of both other classes that may lead the classifier into confusion resulting into compromised performance.

B. FAULTY PREDICTION SCENARIOS

In this section we analyze where the model trained with proposed CaT dataset performs poorly while testing. In Figure 11, we show some visual results and mark the wrongly predicted area with a purple circle in both ground truth images and predicted images. In the top row, we can see the small tree branches laying across the road that are

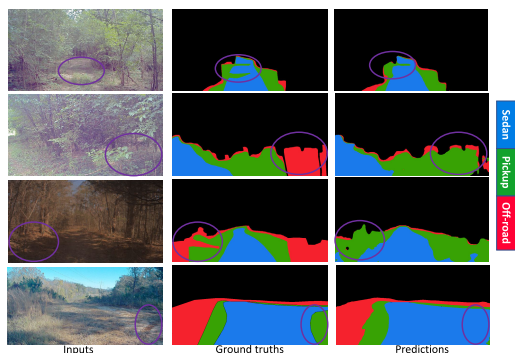


FIGURE 11. Visualizations of some cases where the CaT trained model fails to perform good. Best viewed in color.

labelled as traversable by the pickup-type in ground-truth images but part of the area is predicted to be traversable by sedan-type. Similar types of scenario can be visualized in the third row where a large log is around the edge of the road which is assigned as traversable by the off-road vehicle only. However, it is predicted as traversable by the pickup type. In the image in second row, a densely-bushy area is labeled as traversable by the off-road vehicle only, which is predicted as traversable by the pickup-type vehicle. Furthermore, around the class boundary, we can allocate minor mispredictions. However, the mispredictions are common when there are some confusing area about traversability that happens in case of human-driver too. For example, the bushy areas -as in the image in second row- may look like non-traversable at all or traversable by off-road vehicle, however it may still be traversable by pickup type due to the presence of very weak bushy vegetations.

IX. DISCUSSION

In this section, we explain the basic intuition about traversability in regarding to the model training, its' conceptual similarities and differences with current semantic segmentation aspect, trustworthiness in the output of trained model.

A. TRAVERSABILITY AS A SPECIAL CASE OF SEMANTIC SEGMENTATION FOR AUTONOMOUS DRIVING

Driving in off-road domain is not a straight-forward job like that in on-road. Apart from general scene understanding procedure, maneuvering vehicle's capacity and the nature of the driving trail should be considered while designing autonomous vehicles-based software. Semantic segmentation is a general image/object classification concept in a pixel-wise manner. In case of off-road driving, a range of obstacles/uncertainties across the driving trail are prevalent which then depends upon the maneuvering vehicle's strength. Based on the existing concept of semantic segmentation for road extraction the drivability of the detected trail is ignored. For example, under semantic segmentation concept a log laying across the road could lie in obstacle class no matter how

tall and big is it. Traversability concept can be taken as the special case of semantic segmentation that is specific to the autonomous driving for off-road environment which considers important properties of the objects in driving area from the view point of vehicle's traversing capability. Taking our previous example of log laying across the road, it is possible to classify that part of the scene as a driving trail for the vehicle whose strength is enough to pass over it. Similarly, it is highly probable that, along with the actual track bearing the characters of off-road, some part of the vegetation and forest may come as a traversable path that were to be classified into separate classes under semantic segmentation concept.

B. COMPARING SEMANTIC SEGMENTATION WITH TRAVERSABILITY

In this section, we discuss about the similarities and differences regarding these two paradigms. The semantic segmentation concept is a general aspect of scene-understanding mechanism that classifies each pixel in the image based on the learned properties of the objects in the scene. This concept matches to the traversability concept based on classifying the road-parts corresponding to the vehicle capability. However, their exist the differences based on why traversability of a road is to be determined before driving over it. Through traversability concept, we are trying to establish the relationship between vehicle's driving capability with the nature of the driving path. So it gives the information about which part of the road is difficult by what degree and which vehicle this degree corresponds to, so that the vehicle can traverse it with near-zero damage. On the other hand, semantic segmentation just tells us about which part of the scene belongs to which class -for example: road, forest, obstacle, etc.- that are pre-defined during the process and hence overlooks the correlation between vehicle capability and trail-difficulty. So semantic segmentation concept does not rely on the maneuvering vehicle type, road/trail type, and possibility of passing through. This is considered largely under the traversability concept.

X. LIMITATIONS OF THE STUDY

In this work, we propose the traversability concept for off-road autonomous driving based on the camera images and their labels verified by vegetation overriding and tree density based mathematics using empirical data. Even though the traversability labels are assigned by the experts, there could be inherent flaw in camera images. For example, under the dense vegetation, which looks traversable for an off-road vehicle, there could be a quagmire where a vehicle gets stucked and which thus should be non-traversable. There could be several other such scenarios like hiding an obstacle or a ditch under the tall grass, or the high degree of the soil moisture. From camera images it is hard to tell precisely about such scenarios. So, the traversability assessment only with camera images may not provide the precise information when such scenarios are encountered. Furthermore, we do not consider the possible physical variations like slope changes,

ditches, or open-fields within the dataset. So, a combination of camera images and other sensors like lidar, radar, GPS, etc. is desired to tell about the traversability of an area precisely. For example, the radar returns from a hidden ditch would vary greatly and based on which the corresponding part can be labelled as non-traversable easily.

XI. CONCLUSION AND FUTURE WORK

In this paper, a conceptually new dataset called CAVs Traversability (CaT) dataset is proposed. Due to the nature of off-road autonomous driving, directly using the existing semantic segmentation concept could result in compromised situation. So we propose a traversability concept-based assessment of off-road trail along with CaT dataset considering three different vehicle types. We present different arguments that are inherent with this concept and possibilities of training the models.

As a part of future work, we will be creating the more informative traversability dataset using not only the camera images but with multiple sensors like lidar, radar, and GPS such that the detailed in-depth knowledge of the driving environment could be inferred more accurately and efficiently. Further, we will be considering the effect of hanging tree branches over the driving trail while assessing the traversability as well as demonstrate the training feasibility and the performances for broader spectrum of famous segmentation models with more advance assessment methods.

REFERENCES

- [1] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.
- [2] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018.
- [3] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Jan. 2017.
- [4] L. Dabir, C. Goodin, N. Scherrer, and D. Carruth, "LiDAR data segmentation in off-road environment using convolutional neural networks (CNN)," SAE Tech. Paper 2020-01-0696, 2020.
- [5] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2012, pp. 3354–3361.
- [6] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, *The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results*. Accessed: Jun. 8, 2020. [Online]. Available: <http://www.pascalnetwork.org/challenges/VOC/voc2012/workshop/index.html>
- [7] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 3213–3223.
- [8] G. J. Brostow, J. Fauqueur, and R. Cipolla, "Semantic object classes in video: A high-definition ground truth database," *Pattern Recognit. Lett.*, vol. 30, no. 2, pp. 88–97, 2008.
- [9] P. Sun et al., "Scalability in perception for autonomous driving: Waymo open dataset," 2019, *arXiv:1912.04838*.
- [10] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, "BDD100K: A diverse driving dataset for heterogeneous multitask learning," 2018, *arXiv:1805.04687*.
- [11] A. Valada, G. Oliveira, T. Brox, and W. Burgard, "Deep multispectral semantic scene understanding of forested environments using multimodal fusion," in *Proc. Int. Symp. Exp. Robot. (ISER)*, Tokyo, Japan, Oct. 2016, pp. 465–477.
- [12] P. Jiang, P. Osteen, M. Wigness, and S. Saripalli, "RELLIS-3D dataset: Data, benchmarks and analysis," 2020, *arXiv:2011.12954*.
- [13] S. Sharma, J. E. Ball, B. Tang, D. W. Carruth, M. Doude, and M. A. Islam, "Semantic segmentation with transfer learning for off-road autonomous driving," *Sensors*, vol. 19, no. 11, p. 2577, Jun. 2019.
- [14] D. Wang, "Unsupervised semantic and instance segmentation of forest point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 165, pp. 86–97, Jul. 2020.
- [15] T. Cane and J. Ferryman, "Evaluating deep semantic segmentation networks for object detection in maritime surveillance," in *Proc. 15th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Nov. 2018, pp. 1–6.
- [16] G. Gresenz, J. White, and D. C. Schmidt, "An off-road terrain dataset including images labeled with measures of terrain roughness," in *Proc. IEEE Int. Conf. Auton. Syst. (ICAS)*, Aug. 2021, pp. 1–5.
- [17] S. Hosseinpoor, J. Torresen, M. Mantelli, D. Pitto, M. Kolberg, R. Maffei, and E. Prestes, "Traversability analysis by semantic terrain segmentation for mobile robots," in *Proc. IEEE 17th Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2021, pp. 1407–1413.
- [18] T. Guan, Z. He, D. Manocha, and L. Zhang, "TTM: Terrain traversability mapping for autonomous excavators," 2021, *arXiv:2109.06250*.
- [19] L. Dabir, S. Sharma, C. Goodin, S. Ozier, C. Hudson, D. Carruth, M. Doude, G. Mason, and J. Ball, "Traversability mapping in off-road environment using semantic segmentation," *Proc. SPIE*, vol. 11748, Apr. 2021, Art. no. 117480C.
- [20] *Sekonics SF3325-100 Cameras*, 2017. Accessed: Jun. 2, 2020.
- [21] C. A. Blackmon and D. D. Randolph, "Tr no. 3–783 an analytical model for predicting cross-country vehicle performance; appendix B: Vehicle performance in lateral and longitudinal obstacles (vegetation) volume II: Longitudinal obstacles," US Army Engineer Waterways Exp. Station, Vicksburg, MS, USA, Tech. Rep. 3-783, 1968.
- [22] B. R. Davis, M. D. Neptune, and M. A. Johnson, "Mobility in natural environments; report 1; vegetation override test methods," United States Army Tropic Test Center, New York, NY, USA, Tech. Rep. 7501002, 1974.
- [23] S. A. Shoop, B. Coutermarsh, T. Cary, and H. Howard, "Quantifying vegetation biomass impacts on vehicle mobility," *J. Terramechanics*, vol. 61, pp. 63–76, Oct. 2015.
- [24] G. L. Mason, B. Q. Gates, and V. D. Moore, "Determining forces required to override obstacles for ground vehicles," *J. Terramechanics*, vol. 49, nos. 3–4, pp. 191–196, Jun. 2012.
- [25] M. Rybansky, "Trafficability analysis through vegetation," in *Proc. Int. Conf. Mil. Technol. (ICMT)*, May 2017, pp. 207–210.
- [26] (2016). *All New Teslas are Equipped With NVIDIA's New Drive PX 2 AI Platform for Self-Driving*. Accessed: Jun. 8, 2020. [Online]. Available: <https://www.wall-new-teslas-are-equipped-with-nvidias-new-drive-px-2-ai-platform-forself-driving>
- [27] M. P. Jurkat, P. M. Brady, Jr., and P. W. Haley, "Nato reference mobility model. Edition I. Users guide," Stevens Inst. Technol., Davidson Lab., Hoboken, NJ, USA, Tech. Rep. 12503, 1979, vol. 1.
- [28] *BasicAI*, BasicAI, Irvine, CA, USA, 2020. Accessed: Apr. 24, 2020.
- [29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [30] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2881–2890.
- [31] G. L. Oliveira, W. Burgard, and T. Brox, "Efficient deep models for monocular road segmentation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 4885–4891.
- [32] S. Ruder, "An overview of gradient descent optimization algorithms," 2016, *arXiv:1609.04747*.

• • •