J. Electrical Systems 19-4 (2023): 241-253

[1]Nagula Sai Sathvik
[2] Pudi Manikanta
[3] Kotha Sai Varshith
[4] Pendli Bestha Sai Praneeth
[5] Mohit Lalit
[6] Mohit Angurala

Off-Road Terrain Identification And Analysis

Abstract: - Background: The role of the Terrain is paramount for any autonomous vehicle to drive safely on any type of surface. The Autonomous vehicles should have the capability of identifying the terrain and should adapt to the environment. With the evolution of robotics and Artificial Intelligence, and understanding diverse terrains, the techniques for terrain identification are also advancing with a major focus on safety.

Methodology: To make Terrain Detection and Identification more reliable we used instance segmentation which is a more sophisticated type of segmentation that goes a step ahead of semantic segmentation by performing both object detection and segmentation at the same time. In order to perform Instance segmentation, we used the YOLOv8 architecture which is considered to be the state-of-the-art CNN (Convolutional Neural Network) architecture. The YOLOv8 model was trained on an Off-road Terrain Dataset.

Results: Our findings indicate that the state-of-the-art YOLOv8 instance segmentation model provided the best results for terrain detection and segmentation with a threshold confidence of 0.60, and the results provide a maximum confidence of 0.92 which indicates an accurate segmentation model for the given terrain detection problem.

Conclusion: The present work motivates for a more viable hardware model that makes use of trained computer vision models and cutting-edge sensors that can be tested on different soils and terrain. The results obtained can be used to study about the different Terrains and select the most suitable model, this in turn drives for further research in the subject of Terrain Identification and Detection.

Keywords: Internet-of-Vehicles (IoT), Laser Range Finders (LRF), Accelerometers, Image Segmentation, Image Detection, Instance Segmentation, Semantic Segmentation, Computer Vision, You-Only-Look-Once (YOLOv8), Unmanned Ground Vehicles (UGV), Tensorflow, Pytorch.

I. INTRODUCTION

The study related to Autonomous Vehicles started from the late 1970s when Carnegie Mellon University created its first autonomous technology known as Navlab. Since then, Autonomous vehicles are been widely used in the fields of Transportation, Manufacturing sector, Military, Healthcare industry. The new Generation industries known as Industry 4.0 are the one adopting to new technologies, like IoT, AI, Machine Learning, Big Data, Robotics, Automation, similar thing goes with healthcare 4.0 etc. These new era industries adopting such valuable and reliable technology can enhance the human standard of living.

There has been a significant growth in research related to Autonomous Driving/ Self Driving Vehicles in recent times, many Multi National automobile/technological companies like Mercedes, Nissan, Skoda, Toyota, Google, Rivian, Tesla including many leading technical institutions such as Massachusetts Institute of Technology (MIT), Stanford University and many more have contributed their part in the developing the autonomous industry [1].

The research related to autonomous technology is done in every aspect, from selecting the optimal path to detecting the obstacles in the path. Fig. (1) shows different features required for a fully Autonomous Vehicle to travel from the source to the destination. ,

[1] *Corresponding author: Nagula Sai Sathvik[1*], Chandigarh University, Gharuan, India - 140314.
E-mail *sathvik2210@gmail.com
[2, 3, 4, 5, 6] Chandigarh University, Gharuan, India- 140314

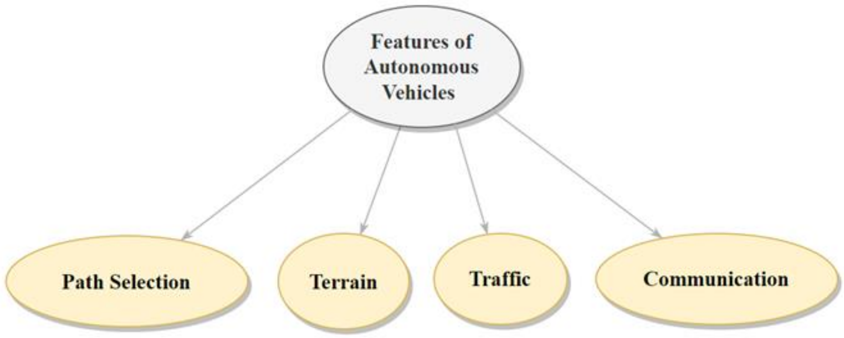

Fig. 1. Features of Autonomous Vehicles

There are many determinants that determine the journey of an autonomous vehicle, these determinants include terrain, traffic, communication, optimum route. The most important factor is the terrain, terrain plays a vital role in determining the speed and acceleration of the vehicle. The driver makes decisions based on the type of terrain and the elevation of the terrain, to control the speed and acceleration of the vehicle. In case of an autonomous vehicle, which have to make the same decisions as a human, based on the type of terrain and elevation requires mixture of hardware and software computation.

The main task is to classify the type of terrain, identify the elevation and adjust the acceleration and speed accordingly. The required features for Terrain Classification and Identification are:

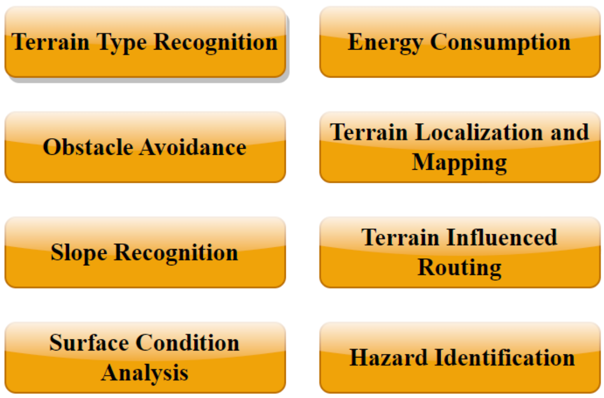

Fig. 2. Features of Terrain Classification and Identification

In order to achieve the required precisions, the autonomous vehicle must be equipped with the advanced sensors, Laser range Finders (LRF), Global Positioning System (GPS), accelerometers, Inertial Measurement Units (IMU), wheel encoders [1]. The collective computation gives the vehicle the smoothest path to go on the most difficult tracks in the world. If the vehicle is able to travel the hardest paths it can easily be able to maneuver the vehicle at any place around the world. The vehicles maneuverability is the most important factor while driving, the steering capabilities of a vehicle depends on the size and dimensions of the vehicle including its tires, body shape and aerodynamics.

As per the literature survey done, the research related to autonomous driving is only been implemented on four wheelers, these four wheelers can be classified based in its dimensions as SUV, Seden, Hatchback, Truck etc., each of it has its own dimensions and steering capabilities. The traversability feature of the terrain is completely different for every type. The dimensions of some of the well-known autonomous vehicles are as follows:

S. No	Vehicle	Type	Institute	Length (inches)	Width (inches)	Height (inches)	Wheelbase (inches)
1	Tesla Model X	SUV	TESLA	198.3	78.7	66	116.7
2	Waymo	SUV	Google	184.3	84.2	61.6	117.7
3	R1T	Pickup Truck	Rivian	217.1	81.8	78.3	135.8
4	IONIQ 5	SUV	Roboride & Hyundai	182.48	74.41	63.98	118.11
5	Mercedes EQS & S class	Seden	Mercedes	207.3	83.7	59.5	126.4
6	Stanley	SUV	Stanford University & Volkswagen	112.4	187.2	75.9	68.0

Table. 1. Feasible Dimensions of Autonomous Vehicles

II. RELATED WORK

Prior to delving into the off-road component of terrain identification, it is imperative that we comprehend the various approaches and procedures utilized in image classification, segmentation tactics, hardware kinds, and deep learning techniques. An approach used in robotics and autonomous navigation systems to determine the type of terrain or ground surface that is near the robot or vehicle is called the “short-range terrain classification based on geometry” algorithm [2]. In order to establish whether the terrain is navigable and drivable for Unmanned Ground Vehicles (UGVs), a terrain traversability analysis is conducted. It highlights open problems and potential directions for mobile robotics, especially in situations where it is unsafe or impractical for people to be present. If a particular vehicle can traverse the trail with little to no damage, we refer to it as traversable for one type of vehicle but traversable for another. Therefore, traversability cannot be only characterized by the characteristics of the driving trail. Additionally, because different vehicle classes have varying traversing capabilities, it depends on the type of corresponding maneuvering vehicle. Terrain traversability refers to the ability of a given vehicle to traverse across a specific type of terrain, such as mud, sand, rocky terrain, snow, and so on. Traversability is therefore the combination of the terrain type, the vehicle type, and the nature of the driving-trail. It describes the ability of a vehicle to traverse across a given terrain with a specific type of trail[3] [4].

The ability of a vehicle to go over various types of terrain, such as mud, sand, rugged terrain, snow, and so on, is referred to as "terrain traversability". Thus, traversability is determined by the mix of the driving trail's features, the kind of vehicle, and the type of terrain. It describes a vehicle's ability to travel on a specific type of trail across a specific type of terrain. Terrain traversability is typically divided into three categories: easy, moderate, and challenging. This classification is based on the amount of energy or effort required for a vehicle to move across a particular terrain. Based on a counting task, Hisham [5] and his team presented their image-level low-count (ILC) supervised density map estimation approach based on an ImageNet pre-trained network backbone (ResNet50) which is built on an ImageNet pre-trained image-level network backbone. Their proposed method has two output sections: image classification and density branch. Du Jiang[6] and their team developed a multiscale target multi-task semantic segmentation model by enhancing the Faster-RCNN model. To expedite dataset preparation, they boosted the training speed by incorporating depth images. Using a Kinect color camera, they captured indoor scene images from diverse angles and backgrounds, constructing an RGB-D dataset for experimental purposes. Instance segmentation presents a significant challenge due to the requirement for precise delineation of object boundaries and handling overlapping objects. In contrast, semantic segmentation proves to be more straightforward as it doesn't necessitate the separation of objects within the same class.

In 2018, Kailun Yang[7] integrated terrain awareness through real-time semantic segmentation. To enhance the effectiveness of segmenting diverse scenes while maintaining efficiency, they crafted an architecture inspired by the Seg-Net-based encoder-decoder, similar to ENet. The decoder architecture incorporated the pyramid pooling segment, drawing inspiration from PSPNet. Additionally, they employed the ADE20K dataset for adaptation, as it encompasses both indoor and outdoor scenarios. Automotive sensors gather information that is processed by the autonomous vehicle's computer to regulate the steering, braking, and speed of the car. Decisions about vehicle control are made using not only the automotive sensors but also data uploaded from other cars and environmental maps stored in the cloud[8]. Without question, LIDARs are the most popular sensors for autonomous cars to track terrain. Instead, there are far more effective radars available. Microwave radar operating at 77 GHz is known as long-range radar. It has a low resolution but can measure speed and identify objects up to 200 metres away.

Sophisticated and reasonably priced, short and medium-range radar operates in the 24 GHz and 76 GHz bands. This sensor can measure distance and velocity, but its resolution is limited by its broad beams and long wavelengths, which also result in complex return signals[9]. In some circumstances, such as inclement weather, radar is more effective than lidar and cameras, but it produces less data and has less angular accuracy than lidar. Radar has lower processing speeds required to handle data output than lidar and cameras, but it does not process any video feeds that contain a lot of data, unlike cameras[10]. Cameras are specialized image sensors that detect the visible light spectrum reflected from objects. Considering how much UV and visible light the sun emits, image sensors are able to pick up a wide range of visible light frequencies. This resembles the way light is perceived by human eyes. High resolution tasks like classification, scene comprehension, and tasks requiring color perception like traffic light or sign recognition are areas in which they excel[11], [12].

When it comes to categorizing road terrain in autonomous land vehicles, the primary mechanism is based on acceleration. This method makes use of information gathered from an accelerometer fixed to the suspension of the car. One of the important procedures in this methodology is determining the experienced vertical acceleration by the vehicle, utilizing a one-quarter vehicle dynamic model estimate road profiles, and extracting different features from this data to allow for the classification of terrain. To be more precise, the accelerometer data is utilized to record the vibration characteristics of the car, which reveal information about the state of the road[12].

Year	Reference	Classification Parameters	Methodology Used	Dataset Used	Research Gaps
2022	[4]	Based on the Dimensions of the vehicle	Semantic Segmentation	CAT Traversability Dataset[4]	Can not distinguish between different instances
2019	[13]	Based on the type of Terrain/Ground	Image Based Terrain Classification – SVM Classifier	Independent Dataset[13]	Accuracy of the sensors used are not feasible for detection in all type of terrains
2019	[14]	Based on the type of Terrain/Ground and Vibrations	DMLPNN (Deep Multi-Layer Perception based Neural Network) Terrain Classifier	Data Collected by Jackal unmanned vehicle, XQ unmanned vehicle[14]	Vibrations are ambiguous at times.
2019	[15]	Based on the type of Terrain/Ground	SVM Classifier	Data collected using Cushman Hauler Pro Cart [15]	SVM Classifier does not gave accurate results in real-time autonomous driving
2022	[16]	Based on the type of Terrain/Ground	ResNet-101 backend pretrained on ImageNet	AI4Mars Dataset[16]	Can not distinguish between different instances
2020	[17]	Based on the type of Terrain/Ground	SCM (Soil Contact Model)	Data Collected from Polaris MRZR 4[17]	The classes are less for real-time road traversal
2021	[18]	Based on the type of Terrain/Ground	SVM Classifier – Semantic Segmentation	Data Collected from ATV-Jr Rover [13]	Results not mentioned
2022	[20]	Based on the type of Terrain/Ground	Negative Obstacle Detector (NODR) and SVM Classifier	Multiple Datasets (CAT[4], Berkeley DeepDive[21] etc)	Review Paper with no actual Experiment Results

Table. 2. Literature Review - (Classification models)

This study offers an overview of how data related to terrain is collected across the regions, the techniques adopted by many researchers for identifying the terrain in different regions. The present work focuses on a much more accurate and reliable terrain identification model mainly focusing on off road terrain.

III. EXPERIMENT

i. Dataset:

The Dataset used for the Experiment to identify the Terrain is Cat: CAVS Traversability Dataset which consists of 3.45k Off-Road Terrain images which were collected by Mississippi State University and is stated in [4]. The data collection was performed near HPCC (High Performance Company Collaboratory) of Mississippi State University. The data is collected from three trails i.e., main trail with a length of 0.64 km , powerline trail with a length of 0.82 km, brownfield trail with a length of 0.21 km . The images were collected considering the required light exposure with different filters compatible with Sekonix SF3325-100 camara model[4].

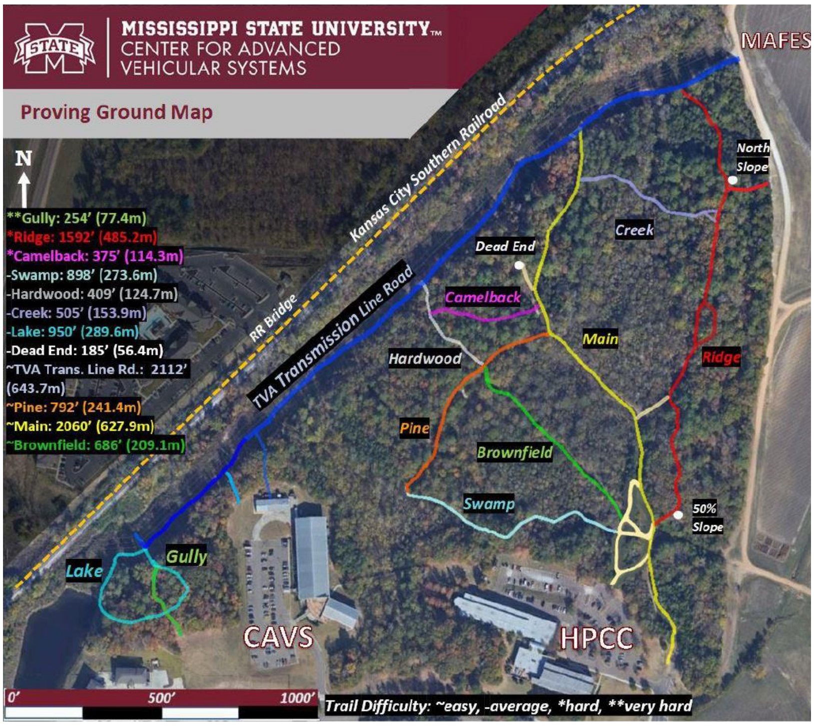

Fig. 3. Data collection Sites of CAT-CavS Traversability Dataset[4]

The data in the dataset CAT-CaVS is used in the present work to identify the Drivable region in an Off-road scenario. The required annotations were performed (The annotations were performed using an Open-source platform Roboflow[13]) on the Dataset in order to label the required Terrian region and segregate the drivable region with the other entities in order to train the model to detect the drivable region in real-time.

There are many other Datasets that can be used for Terrain Identification and Classifications, the dataset must be selected based on the traversability of the vehicle, if the vehicle is to traverse the city roads, suburban roads, the urban/suburban road Terrain Dataset must be selected, if the vehicle have to travel off-road terrain, the dataset according to the region have to be selected, the datasets that are feasible for terrain identification and analysis are:

Dataset	Resource	Region	Purpose
Cat: CAVS Traversability Dataset[4]	Camara	Off-Road	Terrain Identification
Berkeley DeepDrive[14]	Video Sequence	On-Road	Obstacle, Lane Detection
nuScenes Dataset[15]	Camara and Lidar Sensing	On-Road	Object Detection
Open Images V5[16]	Camara	On-Road	Object Detection
Waymo Open Dataset[17]	Video Sequence	On-Road/Off-Road	Environment Awareness
Frieburg Forest[18]	Camara	On-Road/Off Road	Environment Awareness
Oxford Radar RobotCar[19]	Radar Sensing	On-Road	Path Planning

Table. 3. Computer Vision Datasets for Autonomous Driving

ii.Hardware Requirements:

The Hardware Requirements are dependent on different applications of Terrain Identification and Analysis. In the present experiment does not involve any hardware other than the hardware required for the computational task. The hardware configuration used for computation and to identify the feasible/drivable region is:

S No	Device
1	Platform: ASUS Zephyrus G15
2	CPU: AMD Ryzen 9 5900HX 3.30 GHz
3	GPU: NVIDIA GeForce RTX 3050 Ti - 6GB Card
4	RAM: LPDDR5 16 GB Memory

Table. 4. Hardware Requirements

The hardware configuration can vary based on the complexity of the computation; the above given hardware configuration is considered sufficient for training a model with 3.5 k images.

The hardware requirements to control the speed and acceleration of the vehicle based on the slope of the vehicle can also be included in further study of the topic, the required hardware configuration is discussed in the literature part of the paper.

iii. Software Requirements:

The Software Requirements play a vital role in performing the experiment and the required computation. The experiment was performed on python, using different python libraries with suitable and compatible version with the hardware configuration. The Software Configuration include libraries like Ultralytics for importing the required YOLO model, TensorFlow, Pytorch, Keras, Roboflow. The compatible versions adhering the dependencies are:

S No	Library
1	Ultralytics: 8.0.175[20]
2	Tensorflow: 2.10.0
3	Tensorflow-gpu: 2.10.0
4	Pytorch: 2.0.1+cu117
5	Torchvision: 2.0.2+cu117
6	Roboflow: 1.1.7 [13]
7	Keras: 2.10.0
8	Opencv-Python: 4.8.1.78

Table. 5. Software Requirements

The above software configuration was required to perform instance segmentation using YOLOv8 architecture[20].

The Experiment involves the CAT-CaVS Traversability Dataset as discussed earlier in the paper, the dataset was trained using YOLOv8 architecture which is considered to be the state-of-the-art CNN architecture for segmentation.

The Off-Road images in the CAT-CaVS dataset is first annotated using an Open-source annotation platform Roboflow[13]. Out of 3.5k images the most refined images of 1.24 k are selected and the drivable region is labelled for every image. The annotated images are used to train the model with a greater rate of efficiency, as it selects a particular set of pixels to be identified which helps the model to identify the required region in the complete image.

The annotated images are then downloaded into the platform using the required Roboflow program, the annotated images are then used for training the model. The YOLOv8 segmentation model is used in this experiment to perform instance segmentation on the annotated images of CAT-CaVS Dataset.

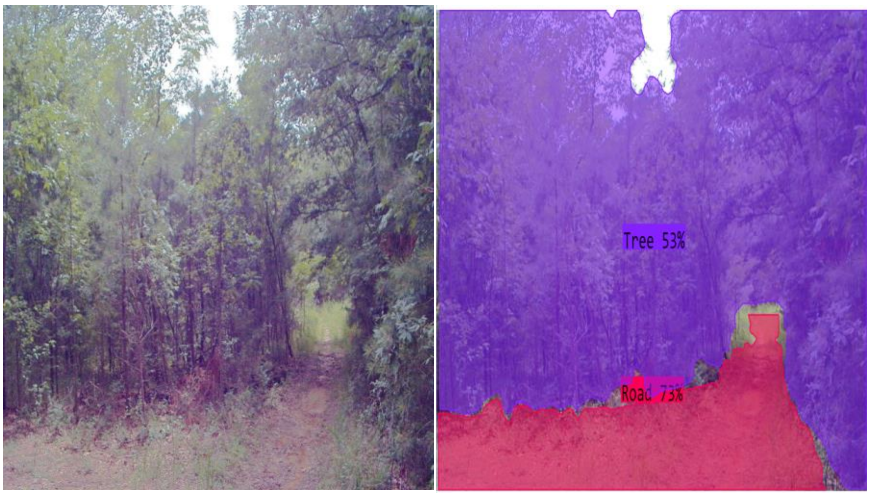

Fig. 4. Raw Image/Annotated Image

Fig. 4 can clearly give an overview of how the image is annotated for training a model[20]. Next comes the Preprocessing stage where the images are reshaped into 640x640 pixels to make it compatible with the YOLOv8 model. The Preprocessing stage is an important step in order to provide accurate and reliable data to the model and makes the computation much simpler.

After Preprocessing the data, Augmentation is performed in order to highlight the colour to the masks provided for segmentation. In this step the mask is given a required colour scheme to differentiate it with the surroundings.

The Pre-processed /transformed data is split into test, train, val(validation) data, the train data consists of 88% images, while the val and test data consists of 8% and 4% respectively. Now the version of the dataset is set to be generated and used for training the model.

The required dataset was loaded into the platform and YOLOv8 segmentation model was trained on the dataset with number of epochs set to 100, image size was set to 640, with batch size set to be 16. The model was trained to segment the feasible drivable region with the other environment [20].

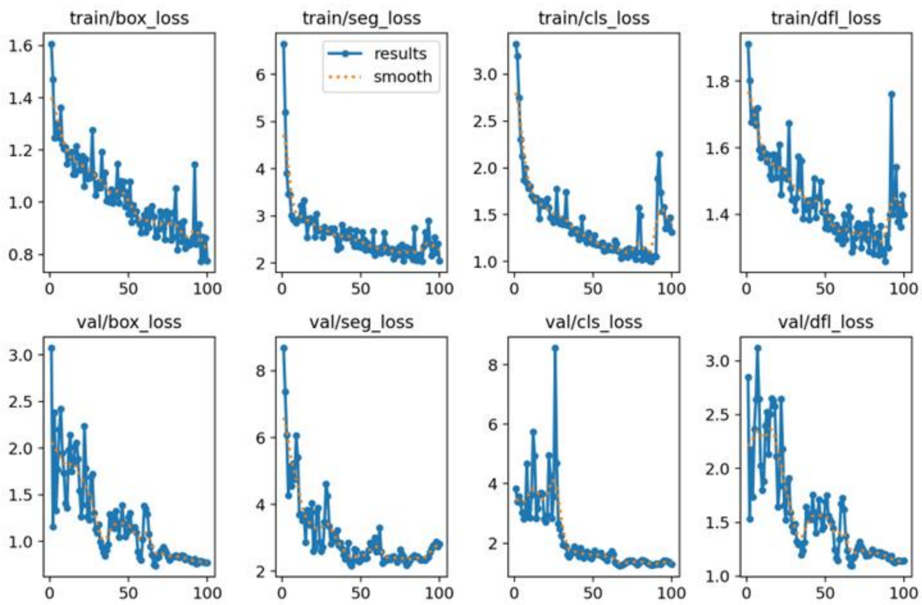

Fig. 5. Loss Functions for Train/Val Data

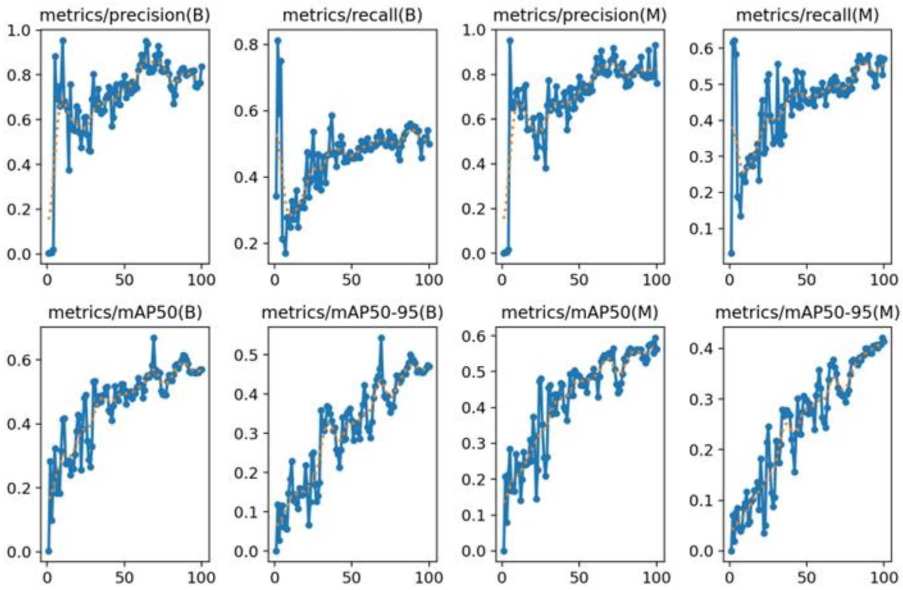

Fig. 6. Training Metrics

The above shown figure illustrates the different loss functions and the performance metrics of the training model. The loss functions include box loss, segmentation loss, class loss for both validation and training data. These loss functions are important to measure the closeness of the predicted segments and boxes with the original. The performance metrics consists of the precision and recall at every stage of training the dataset.

The YOLOv8 model involve feature extraction at multiple stages while the training process, the experiment involved 21 stages of feature extraction while training the dataset. The first and the final stage of feature extraction is as follows:

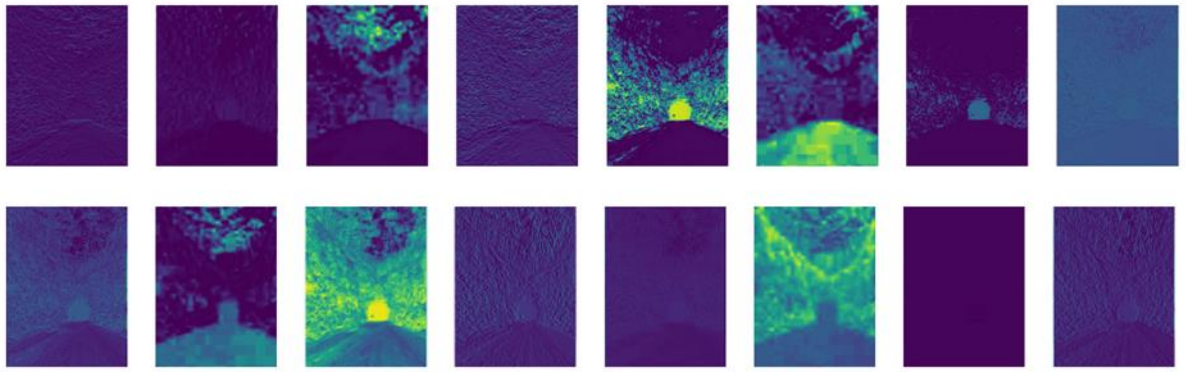

Fig. 7. Stage 1

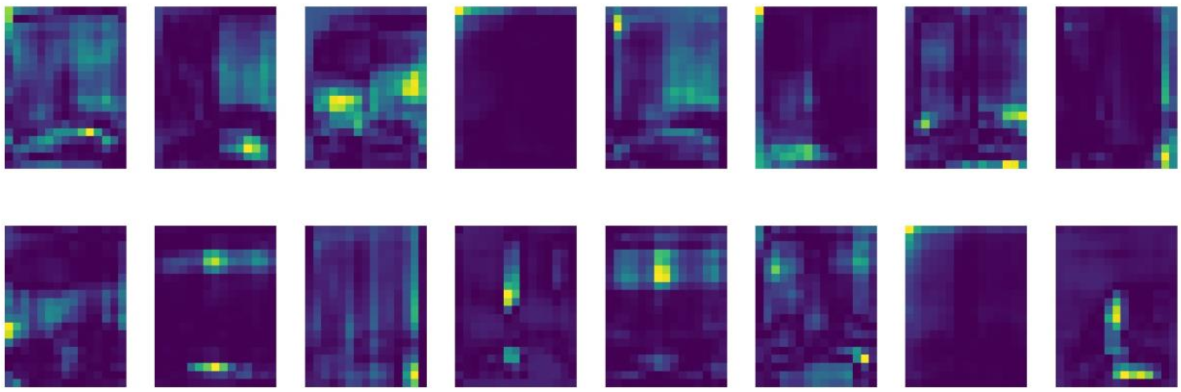

Fig. 8. Stage 21

YOLOv8 Model provides a state-of-the-art segmentation and classification capabilities. YOLOv8 developed by Ultralytics involves a sophisticated structure for segmentation tasks. The YOLOv8 architecture is a combination of multiple Convolutional Layers, Max Pooling, Concat Layers which provide a generalized structure for any segmentation problem[20].

IV. RESULTS

The YOLOv8 model provides a sophisticated instance segmentation model for image identification and analysis. Instance segmentation model is considered to be more finely grained and refined than semantic segmentation model to provide a detailed view of different instances in a road terrain scenario.

The experiments performed in, is a semantic segmentation model which only provides a bigger picture of the terrain identified, but instance segmentation provides a more refined identification of the terrain, even detecting obstacles and trees in an Off-Road scenario. The results of the experiment conducted in the present work finely segments the tress, small plants, drivable road, obstacles etc, while in the required road is only identified while neglecting the other areas of the image.

Instance Segmentation model is useful to identify and classify between different instances of the image giving a clearer view of the other objects and increases the awareness score of the traversal. The trained instance segmentation model gave some accurate results as discussed in the experiment section.

The Trained model is then subjected to validation in order to check the closeness of the feature extracted with that of the trained data with that of the features extracted with the validation data. The trained model gave the following results while the validation stage of the experiment:

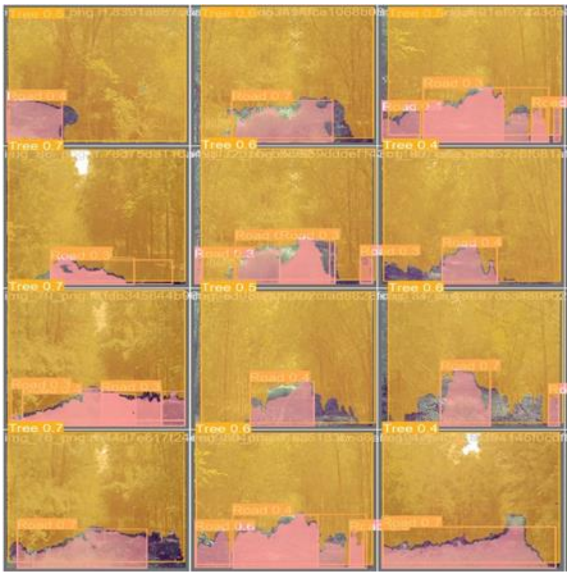

Fig. 9. Train Results

The model is then prepared for further tracking and predictions. The model after validation an image is taken for predictions and detection of the required drivable region. The desired drivable region is detected from the image with the help of the trained Instance Segmentation model, which clearly bifurcated the drivable region with the other environment. The predicted images are as follows:

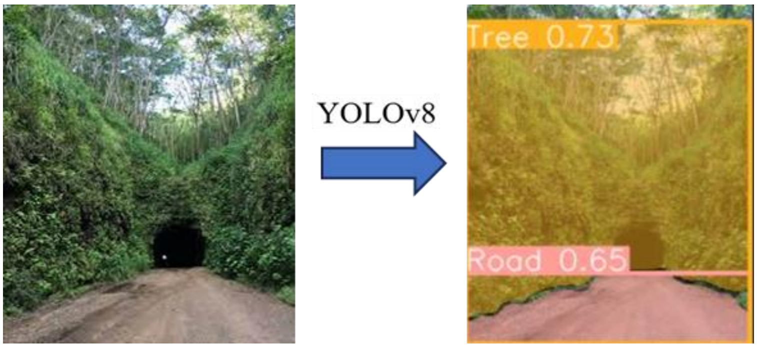

Fig. 10. Segmentation Result-1

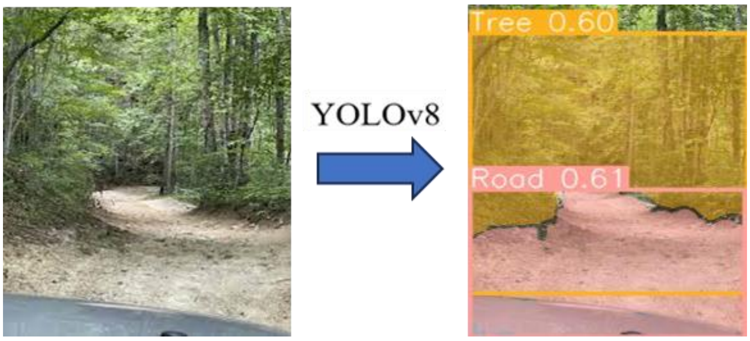

Fig. 11. Segmentation Result-2

The model was then tested based on different performance metrics which include F1 Score, Precision, Confidence, etc. The Curves for the performance metrics are as follows:

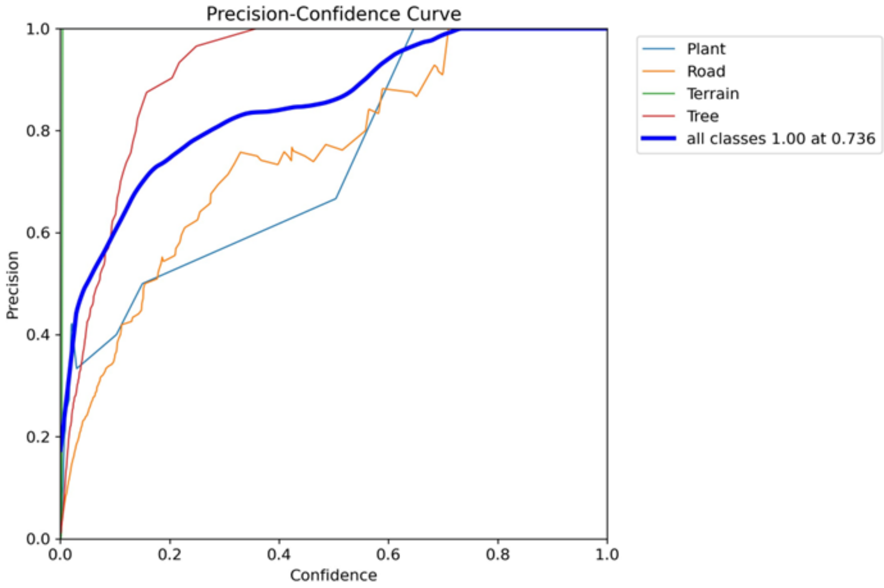

Fig. 12. Precision- Confidence Curve

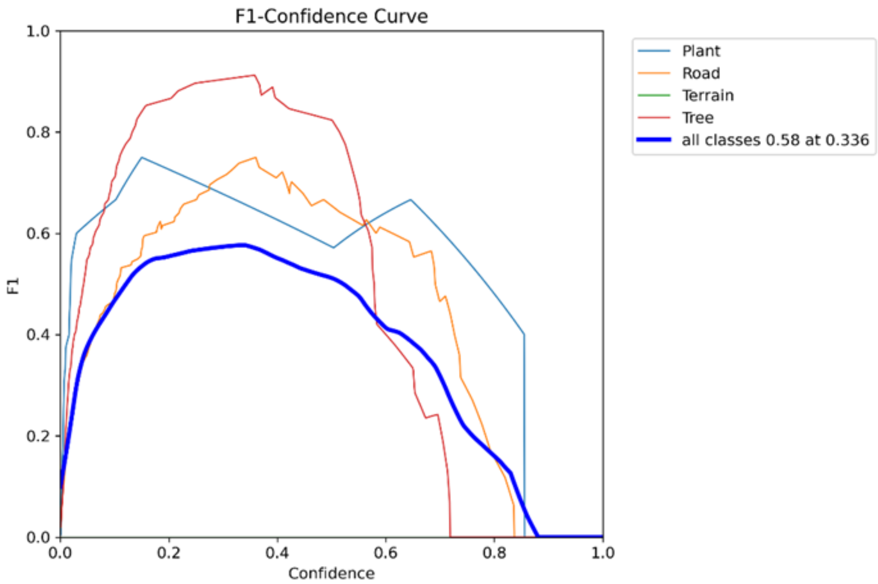

Fig. 13. F1-Confidence Curve

V. CONCLUSION AND FUTURE SCOPE

Over the years the industry of Autonomous Vehicles has seen a rapid growth and the growth in the future is also imminent. The OEM's (Original Equipment Manufacturer) are adopting to new technologies like IoT, Artificial Intelligence, Machine Learning while increasing their R&D capabilities, this motivates for further research in the field of autonomous driving.

The research provides an answer to "How can a terrain or a drivable region be detected by an autonomous vehicle?". The implementation in the present work only focuses on the computational research, which motivates to further continue the research in terms of a more practical hardware implementation.

The present work can provide computational capabilities to any hardware working model and can be used to perform terrain identification in real-time. The Future work for the field terrain identification will include a working hardware model that can provide a much more generalized and sophisticated model that can identify any type of terrain and can control the maneuverability of the vehicle accordingly.

The model is aimed to be trained on a more accurate dataset collected with advanced camara modules working in real-time. The existing terrain model can further be integrated with models that can be used for path planning, object detection, GPS; aiming to build a fully autonomous vehicle in the future.

REFERENCES

[1] S. Wang, "Road Terrain Classification Technology for Autonomous Vehicle." [Online]. Available: http://www.springer.com/series/15608

[2] M. Bajracharya, A. Howard, L. H. Matthies, B. Tang, and M. Turmon, "Autonomous off-road navigation with end-to-end learning for the LAGR program," *J Field Robot*, vol. 26, no. 1, pp. 3–25, Jan. 2009, doi: 10.1002/rob.20269.

[3] P. Papadakis, "Terrain traversability analysis methods for unmanned ground vehicles: A survey," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 4. Elsevier Ltd, pp. 1373–1385, 2013. doi: 10.1016/j.engappai.2013.01.006.

[4] S. Sharma *et al.*, "CaT: CAVS Traversability Dataset for Off-Road Autonomous Driving," *IEEE Access*, vol. 10, pp. 24759–24768, 2022, doi: 10.1109/ACCESS.2022.3154419.

[5] H. Cholakkal, G. Sun, F. Shahbaz Khan, and L. Shao, "Object Counting and Instance Segmentation with Image-level Supervision."

[6] D. Jiang, G. Li, C. Tan, L. Huang, Y. Sun, and J. Kong, "Semantic segmentation for multiscale target based on object recognition using the improved Faster-RCNN model," *Future Generation Computer Systems*, vol. 123, pp. 94–104, Oct. 2021, doi: 10.1016/j.future.2021.04.019.

[7] *2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2018.

[8] S. M. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive Radars: A review of signal processing techniques," *IEEE Signal Process Mag*, vol. 34, no. 2, pp. 22–35, Mar. 2017, doi: 10.1109/MSP.2016.2628914.

[9] S. Royo and M. Ballesta-Garcia, "An overview of lidar imaging systems for autonomous vehicles," *Applied Sciences (Switzerland)*, vol. 9, no. 19, Oct. 2019, doi: 10.3390/app9194093.

[10] Telecommunications Society, Univerzitet u Beogradu. School of Electrical Engineering, IEEE Communications Society. Serbia & Montenegro Chapter, and Institute of Electrical and Electronics Engineers., *TELFOR 2018 : 26. Telekomunikacioni Forum : Beograd, 21 i 22. novembar 2018. godine, Sava T͡sentar = TELFOR 2018 : 26th Telecommunications Forum : Belgrade, 21 and 22 November 2017, the SAVA Center*.

Engineers., *TELFOR 2018 : 26. Telekomunikacioni Forum : Beograd, 21 i 22. novembar 2018. godine, Sava T͡sentar = TELFOR 2018 : 26th Telecommunications Forum : Belgrade, 21 and 22 November 2017, the SAVA Center.*

[11] J. Steinbaeck, C. Steger, G. Holweg, and N. Druml, "Next generation radar sensors in automotive sensor fusion systems," in *2017 Symposium on Sensor Data Fusion: Trends, Solutions, Applications, SDF 2017*, Institute of Electrical and Electronics Engineers Inc., Dec. 2017, pp. 1–6. doi: 10.1109/SDF.2017.8126389.

[12] Institute of Electrical and Electronics Engineers, *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring) : proceedings : Nanjing, China, 15-18 May 2016*.

[13] V. van der Burg, G. de Boer, A. A. Akdag Salah, S. Chandrasegaran, and P. Lloyd, "Objective Portrait," Association for Computing Machinery (ACM), Jul. 2023, pp. 387–400. doi: 10.1145/3563657.3595974.

[14] H. Xu, Y. Gao, F. Yu, and T. Darrell, "End-to-end Learning of Driving Models from Large-scale Video Datasets," Dec. 2016, [Online]. Available: http://arxiv.org/abs/1612.01079

[15] H. Caesar *et al.*, "nuScenes: A multimodal dataset for autonomous driving," Mar. 2019, [Online]. Available: http://arxiv.org/abs/1903.11027

[16] I. Krylov, S. Nosov, and V. Sovrasov, "Open Images V5 Text Annotation and Yet Another Mask Text Spotter," Jun. 2021, [Online]. Available: http://arxiv.org/abs/2106.12326

[17] P. Sun *et al.*, "Scalability in Perception for Autonomous Driving: Waymo Open Dataset," Dec. 2019, [Online]. Available: http://arxiv.org/abs/1912.04838

[18] A. Valada, R. Mohan, and W. Burgard, "Self-Supervised Model Adaptation for Multimodal Semantic Segmentation," Aug. 2018, doi: 10.1007/s11263-019-01188-y.

[19] W. Maddern, G. Pascoe, C. Linegar, and P. Newman, "Year, 1000km: The Oxford RobotCar Dataset." [Online]. Available: http://robotcar-dataset.robots.ox.ac.uk

[20] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection." [Online]. Available: http://pjreddie.com/yolo/

[21] Nagesh Appe, S. R. ., Arulselvi, G. ., & Balaji, G. . (2023). Tomato Ripeness Detection and Classification using VGG based CNN Models. International Journal of Intelligent Systems and Applications in Engineering, 11(1), 296–302. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2538