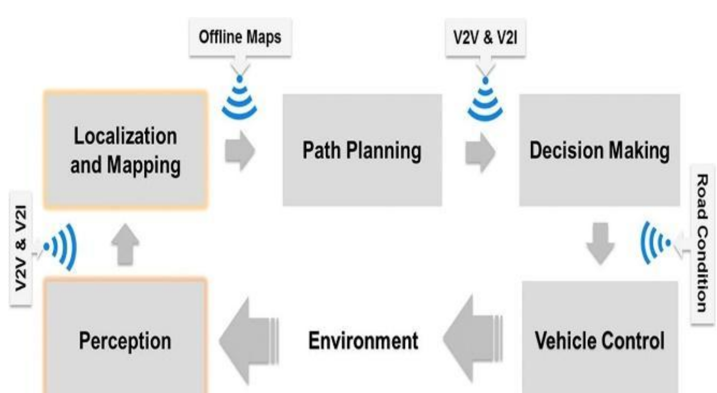
Perception, Localization, and Mapping using deep learning

INTRODUCTION:

Deep learning with its cardinal contribution to artificial intelligence has brought a new dawn to the field. As we dive into the concept of Deep learning topics like computer perception begins to kick in and thus begin the quest to achieve human reflexes, recognition, and comprehensive analysis. Perception being one of the few attributes of this project has enabled us to learn the concept of not only Autonomous vehicles but also computer vision techniques, robot controller programming, semantic segmentation, object detection, and localization. The research aims to establish a methodology to detect objects using the Yolov7 algorithm. The "You Look Only Once" can be used for instances where the identification of certain objects is not distinct. Hence, perception depends on a computer's ability to perceive things. The same can be developed or enhanced by feeding data to the system i.e., its training and testing with proper classifiers. As the vehicle follows a definite path an autonomous vehicle must distinguish between objects and their proximity to the vehicle. In the past 20 years, Yolo and its variants have been dominating the field of object detection and have given a significant increase in the accuracy of performance output. As the vehicle runs on the road the real-time image processing and object detection should be precise, also having time to deal with critical situations and meanwhile performing those computations. is a tedious task.



The stream of autonomous vehicles is not only limited to vehicles on the ground, it has its branches towards the moon with rockets and all the way down to the exploration challenge deep with automated submarines. The analysis of places that are unreachable to humans. The introduction of an embedded system has established new ways of vehicle-to-environment interaction. Unfortunately, this report is a small ramification of localization of robots on a road.

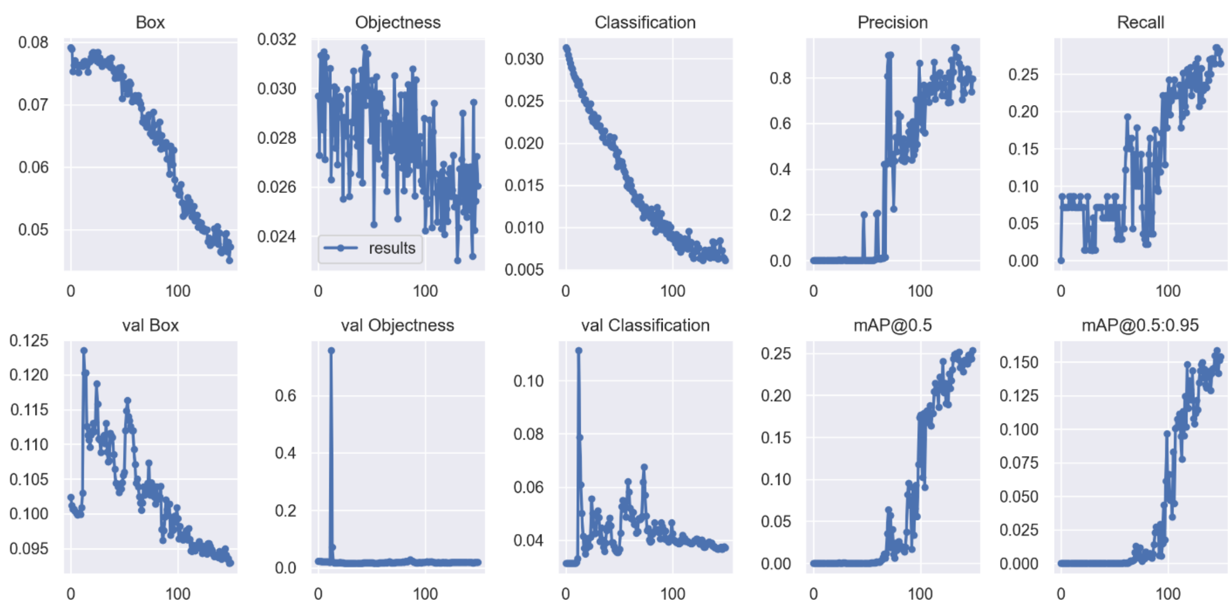
CONCEPTS COVERED:

The cardinals demonstrated a real-time simulation of object detection and semantic segmentation in an autonomous vehicle. While moving along straight and curved roads the speed and position can be critical to the object. Thus, having proper image classifiers along with a strongly trained model only can achieve precise results. The Tesla car testing is an example of such a critical instance where the vehicle couldn't identify the stroller and run over it. The deep learning model prepared for the project provides an upper hand over the machine learning model because it is better at recognizing curved paths. Nevertheless, the outputs generated through neural networks are Blackbox, it can have both better and weaker comprehension of the classes provided to it even with accurate labeling. Furthermore, both color-based and shape-based features are complementary to each other. Capturing them gives a brief idea about the physical attributes of the object. The role played by semantic segmentation is representing each pixel to its corresponding class called dense prediction. The segmentation is not instantaneous and hence has a different model for instant segmentation. This is because it does not decipher between two equivalent input images of a particular object. Collectively everything works simultaneously to maintain stable and concise movement. The tools used for the process are labeling for labeling images, PyTorch library, and the Yolo v7 algorithm. Alternatively, we could compute the correlation between the car image and the test region. Either way, this general approach is known as template matching. Furthermore, while marking the classifiers the elimination of classes with an inaccurate view should be eliminated to avoid confusion for the system.

PROPOSED WORK:

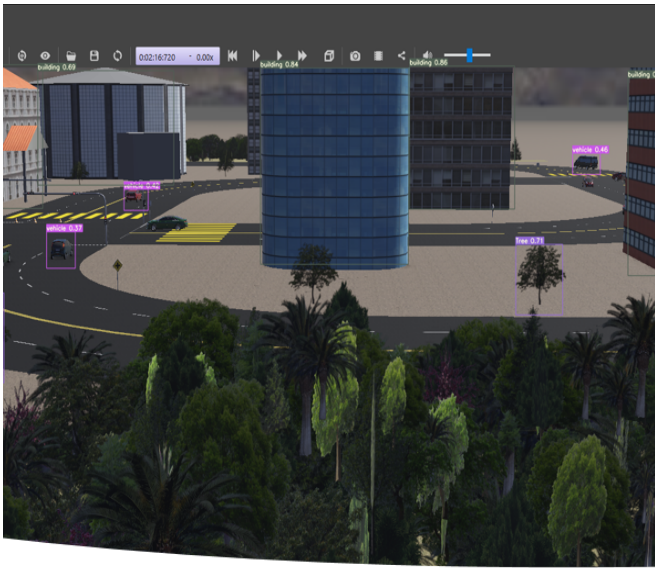
The simulation of the project is done on Webots R2022a. The platform provides nodes(objects) to create an arena in which smaller nodes are used to create a world. It provides geometrical nodes as we contrast the desired robot. However, maps can be downloaded using both SUMO and JOSM. In practice, Java OpenStreetMap gives less efficiency in detail than SUMO. The object detection done by the YOLO V7 algorithm is the training of the model with thousands of datasets to get an accuracy of 0.5 in the training. The training had a lot of challenges.Moreover we have used the open source tool MMSegmentation which is built for custom training , inference , developing various segmentation models . One of the main challenges was to train the model perfectly so that it perceives edges and identifies as it was getting confused between various look-alike classes such as trees and pedestrians, pole, and tree, etc. The data sets used for marking these images are the KITTI dataset which had over 1000 images with vehicles, traffic signs, traffic lights, and pedestrians for the labeling process. There are some more data sets used to augment the identification process and many pre-labeled data sets for further training and testing. The robot has basic characteristics like hinge joints and a motor which are saved as a .msg file and can be accessed through the name of the controller in the Webots simulator. The simulator needs a powerful GPU and for this project, we have used NVIDIA RTX for faster simulation of the vehicle.

RESULTS

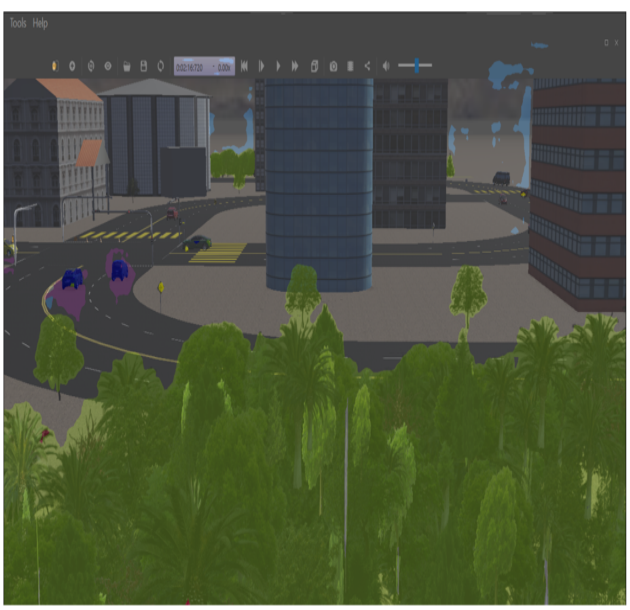


results of training yolov7 custom data

OBJECT DETECTION



SEMANTIC SEGMENTATION



FUTURE WORK:

In the future, we will try to implement localization using other alternatives and train with more classifiers. Moreover, the model can be further improved to have embedded systems and better perception with emergency crash calling, and vehicle-to-vehicle communication and ensure newer safety protocols. The knowledge acquired by this project will come in handy while developing vehicles subjecting to a 3D orientation such as drones and submarines. The development of newer algorithms is also one of the major contributions regarding lane and object detection to solve image processing issues that arise in situations of dense fog or darkness. Implementation of object detection can be done for the systems not only vehicles such as the detection of facial features, and recognition of moving vehicles using DeepCNN-SLAM. Therefore, the topic has a lot of ramifications to it and has a vast playground to play with this latest technology.

CONCLUSION:

The successful representation of objects according to their class can be seen in the project. We couldn't accomplish proper localization and mapping, but real-time object detection was achieved. The classifiers were corroborated as labeled and a fully working simulation with real-time is obtained. The testing and training of the model have given us the essence of deep learning. The understanding of the perception through object detection and semantic segmentation of a vehicle is comprehended thoroughly through the project.

REFERENCES:

- Autonomous Driving with Deep Learning: A Survey of State-of-Art Technologies
- YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors
- Semantic Segmentation for Autonomous Driving: Model Evaluation, Dataset Generation, Perspective Comparison, and Real-Time Capability
- YOLACT: Real-time Instance Segmentation
- Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges

- Kitti dataset , Cityscapes , roboflow
- Training yolov7 custom using custom data
- labeling Images Tool
- open source segmentation tool –MMSegmentation