Semantic Segmentation for a Self Driving Vehicle

WPI Robotics Engineering - RBE549 Term Project

Zeke Flaton
WPI Robotics Engineering Dept
Worcester Polytechnic Institute

Keith Chester
WPI Robotics Engineering Dept
Worcester Polytechnic Institute

Robert DeMont
WPI Robotics Engineering Dept
Worcester Polytechnic Institute

***Abstract*—Semantic segmentation for a self driving vehicle achieved in live simulation using CARLA. A deep neural network will be utilized to categorize pixels in an image generated by a virtual front-facing camera on our vehicle.**

I. Introduction

For our term project, we plan to build a system that can perform semantic segmentation on images with a focus on self driving vehicles. Our goal will be to distinguish different parts of an image so that the self driving vehicle can distinguish other vehicles and pedestrians. Our approach will incorporate training a deep neural network as the basis for fast, efficient, correct image segmentation.

II. Project Goal

We intend to create a deep neural network to detect and segment images of common scenarios encountered during normal driving. We plan to implement this in the CARLA simulator to test against realistic activity. We anticipate that our image categories will be correctly labelled via colors to highlight the segmentation into designated categories. We will track picture level accuracy of objects in the scene while running a simulation in CARLA showing synchronized semantic segmentation.

III. Problem Statement

A. Previous Work

As noted in [1] many previous efforts have accumulated human labeled datasets of various sizes to enable deep learning for image segmentation. This is time-consuming, costly, and tedious. For expediency, most research is now incorporating use of video games scenarios and simulations to quickly advance their work [2] and [3].

B. Significance

Today most automakers are endeavoring to incorporate some level of autonomy in their offerings while the general public is alternately enthusiastic and hesitant to accept and adopt these systems. In the US alone, there were over 12MM traffic crashes in 2019 involving nearly 36,000 fatalities. The NHTSA estimates that 94% of these were caused by human errors. The NHTSA believes autonomy will greatly reduce these numbers as well as have desirable effects on efficiency, convenience, economic, and environmental concerns and has advanced policies to encourage various levels of autonomy for road use.

Given that auto manufacturers are using some version of image segmentation in autonomous commercial products already, we are unlikely to add significant contributions to the general knowledge base. However, for our continued progress in the field of robotics, a practical, working knowledge of image segmentation will be valuable in future projects including our capstone exercises.

IV. Methodology

For this project we aim to utilize CARLA, a simulator developed by Intel to assist the research and development of technologies for autonomous vehicles. The simulator creates a virtual neighborhood with vegetation, traffic signs, other moving vehicles, and pedestrians. In this simulator we will have a target car with a virtual front-facing camera.

The simulator provides an additional ”view” for this camera; since it has perfect knowledge of the location and category of all objects within the simulation, it can correctly identify and categorize all pixels witnessed by the front-facing camera. CARLA provides a dozen categories such as vehicles, pedestrians, buildings, fences, traffic signs, vegetation, and so on. To limit the scope of the project due to limited time we shall be restricting our efforts to two important categories to be found on the roads - vehicles and pedestrians.

To further aid our limited timeline, we are augmenting our available data beyond our usage of CARLA. Since the public release of the tool numerous datasets of the segmentation camera exist. Since deep learning model performance is directly proportional to the availability and depth of data available, we will be utilizing any publicly available datasets to augment our own work.

To build, train, and execute the deep learning models for semantic segmentation we will be utilizing the Tensorflow framework and platform, created by Google. We choose this framework for its vast popularity and expressiveness for training neural networks.

To train the neural network, we will need to utilize GPUs. While some members of our team have GPUs available for local training, not every team member does. These team members will utilize Google Colab, a product providing free hosted Jupyter Notebooks with the option to request free or inexpensive rented GPUs for Python environments. To provide datasets to the notebook, Google Drive, a networked file storage service, can provide the notebooks with access to the datasets and provide a storage solution for training checkpoints.

Once trained, we will create a script that will run the CARLA simulator and output our model's resulting segmentation map over the front-facing camera's output, to provide an easily accessible demonstration of our model's accuracy.

A. Challenges

We suspect we will encounter several challenges during the course of this project. We discuss some of them here.

The balance of labels within our dataset will provide a significant challenge. It is likely that the presence of vehicles will vastly exceed the presence of pedestrians within our dataset, leading to an imbalance that can lead to improper labeling of our subjects. We may have to create additional scripts to filter data from becoming too one-sided during training, but this will further reduce the size of our dataset.

We are aware that training a semantic segmentation model can take a significant amount of time and data. While the existence of simulators such as CARLA greatly assist with this - we can always generate *more* data - this would take substantial investments in time. We hope that the utilization of publicly accessible datasets already generated with CARLA will ease this challenge.

GPU utilization is required for training a model of this complexity and size, and GPU ownership in our group is limited. While we have proposed Google Colab as a solution to this, it creates additional problems. Connecting Google Drive as a network drive for dataset access and checkpoint saving is significantly slower than a local disk, and will increase training time greatly. The usage of GPUs within Colab can also be bound to time limits and resource availability, so we can not guarantee the reservation of a set time or size of GPU. This will routinely interrupt attempts at training models.

B. Timeline

We propose the following timeline for this project:

Task	Date
10/21	Establish CARLA + training environments
10/28	Decide on model network architecture
10/28	Identify existing datasets
10/28	Develop scripts for generating more via CARLA
11/4	Train model
11/11	Working Sim + Model
11/18	Results tabulation
12/2	Report writing and Presentation development
12/9	Demo and Report Turn-in

TABLE I
PROPOSED TASK TIMELINE

V. METRICS

Success will be measured by:
1. A working simulation generating a real-time segmented image interlaced with the real world image from CARLA
2. 90+% accurate segmentation of vehicles and pedestrian objects.

VI. SUMMARY

From this project, we intend to learn about the application of image segmentation in autonomous vehicle navigation and how it might be used for autonomous vehicle control input. At this point we expect equal contributions by team members on designing methods, programming solutions, debugging outcomes, and reporting results.

REFERENCES

[1] Liu, X., Deng, Z. & Yang, Y. Recent progress in semantic image segmentation. Artif Intell Rev 52, 1089–1106 (2019). https://doi.org/10.1007/s10462-018-9641-3
[2] S. Malec, 'Semantic Segmentation with Carla Simulator', Dissertation, 2021.
[3] D. R. Niranjan, B. C. VinayKarthik and Mohana, "Deep Learning based Object Detection Model for Autonomous Driving Research using CARLA Simulator," 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021, pp. 1251-1258, doi: 10.1109/ICOSEC51865.2021.9591747.
[4] C. Wang, '2D object detection and semantic segmentation in the Carla simulator', Dissertation, 2020.