Planning Report - Deep Machine Learning

Road Image Segmentation for AD/ADAS related applications

(Subject Code: SSY340)

Project Planning Group 27

Adithyaa Ramesh
Dimas Rizky Kurniasalim
Dion Lim
Mathanesh Vellingiri Ramasamy

CHALMERS
UNIVERSITY OF TECHNOLOGY

Submitted on Wednesday 9th October, 2024

Problem Statement and Motivation

Road Image Segmentation is a key task in autonomous driving (AD), advanced driver assistance systems (ADAS), traffic monitoring, and smart city development applications. It helps us to understand the driving environment which leads to safer and more efficient decisions. This is especially important for applications like lane detection, obstacle avoidance, and path planning in autonomous vehicles. Here, we plan to investigate the challenge of developing and improving neural network models to perform semantic road segmentation and compare the performance of different architectures. Specifically, we aim to explore how the different model architectures, (e.g., fully convolutional networks, U-Net, and SegNet) perform on different datasets (Cityscapes, nuScenes, KITTI). Our results will offer insights into model effectiveness in terms of accuracy and cross-dataset performance.

Accurate and reliable road segmentation in a potentially crowded image is essential for ensuring safety in autonomous vehicles, optimizing traffic flow, and enhancing urban planning, especially in busy streets. The findings will help improve the effectiveness of segmentation models in real-world autonomous driving applications, enhancing safety and reliability.

We also aim to observe insights into which architectural features (e.g., depth of the network, use of skip connections, attention mechanisms) contribute to the best segmentation performance. Finally, by comparing against different datasets, we can also evaluate how well the models generalize across different environments (under challenging conditions like occlusions, poor weather, or varying road textures).

Data Sets

We plan to use multiple datasets to enhance the robustness of our models through cross-validation. In particular, the below-listed datasets are to be used:

- KITTI Road Dataset
- Cityscapes Dataset
- nuImages Dataset from nuScenes

These datasets provide diverse urban driving scenarios, allowing us to train and test our models under various conditions and compare the performance and accuracy across datasets. We could also consider using other datasets like Lyft, SYNTHIA, and Oxford Road Boundaries based on their availability and annotation data.

Pre-trained Models

Pre-trained models for SegNet, UNet, and VGG Net are available in public repositories, which we will use as baselines. However, we plan to create our own fully convolutional network and also modify the VGG-based model architecture for road segmentation, and cross-validate the same across different datasets, which introduces a novel aspect to this study. In particular, we will implement the modifications, transfer learning, and training process ourselves to allow comparisons and possibly introduce improvements, such as integrating additional classes for segmentation if time permits.

Evaluation

The models will be evaluated using **Intersection over Union (IoU)** and **pixel-wise accuracy metrics** for segmentation. **Cross-validation** will involve training on one dataset and testing on another to measure the generalization capability of our models. Finally, we can test our model across **various weather and lighting conditions**, **different road textures** found in the datasets to see the effect of generalization. We could also evaluate **computational efficiency** (real-time performance), as it is a crucial consideration for AD/ADAS systems. Some of the other criteria that we can look up are the model's capability in handling **edge cases and failures** (e.g., misclassifications near occlusions, shadows, or cluttered environments). **Dice Coefficient** can be used as well (measuring the similarity between predicted and ground-truth segments, balancing precision and recall).

Time Plan

This is a rough estimation of how our workflow will look like for this project.

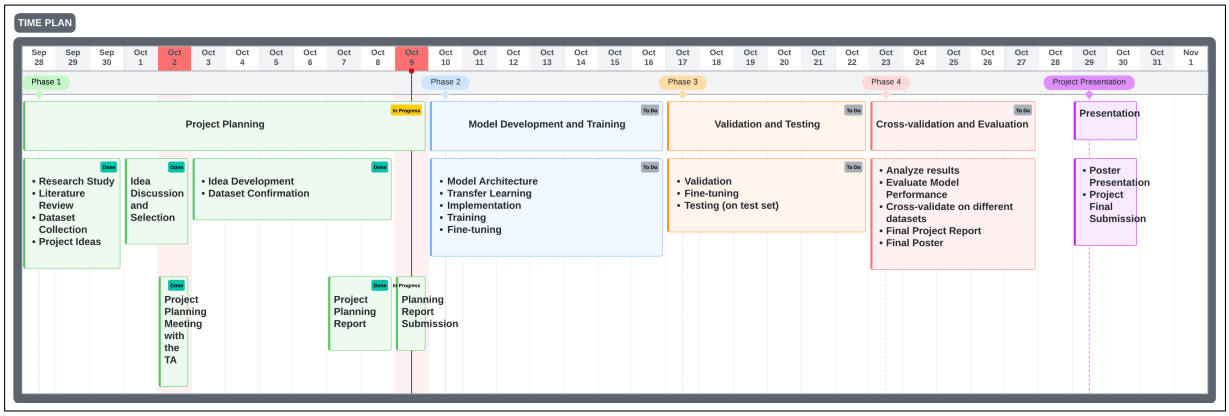

Figure 1: *Project Time Plan*

References

Some of the relevant papers that we study for this project are as follows:

- nuScenes: A multimodal dataset for autonomous driving. (Authors: Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, Oscar Beijbom)
- Very Deep Convolutional Networks for Large-Scale Image Recognition (10 Apr 2015). (Authors: Karen Simonyan, Andrew Zisserman)
- SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. (Authors: Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, Ping Luo)
- RoadSegNet: a deep learning framework for autonomous urban road detection. (Authors: Kushagra Pal, Piyush Yadav and Nitish Katal)
- Intelligent Semantic Segmentation for Self-Driving Vehicles Using Deep Learning. (Authors: Qusay Sellat, SukantKishoro Bisoy, Rojalina Priyadarshini, Ankit Vidyarthi, Sandeep Kautish, Rabindra K. Barik)
- DeepAerialMapper: Deep Learning-based Semi-automatic HD Map Creation for Highly Automated Vehicles. (Authors: Robert Krajewski, Huijo Kim)
- Lidar Panoptic Segmentation in an Open World. (Authors: Anirudh S Chakravarthy, Meghana Reddy Ganesina, Peiyun Hu, Laura Leal-Taixe, Shu Kong, Deva Ramanan, Aljosa Osep)
- Task-Oriented Pre-Training for Drivable Area Detection. (Authors: Fulong Ma, Guoyang Zhao, Weiqing Qi, Ming Liu, Jun Ma)
- U-Net: Convolutional Networks for Biomedical Image Segmentation. (Authors: Olaf Ronneberger, Philipp Fischer, and Thomas Brox)
- Image Segmentation Using Deep Learning: A Survey. (Authors: Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, Demetri Terzopoulos)
- Techniques and Challenges of Image Segmentation: A Review. (Authors: Ying Yu, Chunping Wang, Qiang Fu, Mingliang Gao, et al.)
- Fully Convolutional Networks for Semantic Segmentation. (Authors: Jonathan Long, Evan Shelhamer, Trevor Darrell)