

DATS 6303: Deep Learning

Final Project- Proposal

Title: High-Performance Semantic Segmentation for Autonomous Urban Navigation

Purpose:

While semantic segmentation of urban scenes is a well-established area in computer vision, this project offers an opportunity to deepen our understanding of its practical applications and challenges, especially within the context of autonomous driving. Through the implementation and analysis of segmentation techniques, we aim to explore how fine-grained, pixel-level scene understanding contributes to the broader goal of safe and intelligent vehicle autonomy.

This work allows us to examine how segmentation models can differentiate between critical classes—such as roads, vehicles, pedestrians, and urban infrastructure—and how these classifications feed into downstream tasks like navigation, obstacle avoidance, and real-time decision-making. By evaluating segmentation accuracy, class imbalance, and model robustness in real-world urban conditions, we gain insight into both the technical limitations and potential enhancements for future systems.

Dataset:

We will use the A2D2 (Audi Autonomous Driving Dataset), which provides sensor data from Audi vehicles designed for autonomous driving research:

- We'll use approximately 15,000 frames with semantic segmentation annotations (a substantial subset of the full dataset). This subset is sufficient for training a DeepLabV3+ model, especially with:
 - ❖ Data augmentation (flipping, cropping, scaling)
 - ❖ Transfer learning (e.g., using a pretrained ResNet-50 backbone)
- We'll work with images at 1280×720 resolution to balance quality and computational efficiency

Dataset URL: <https://www.a2d2.audi/a2d2/en/dataset.html>

Deep Network Architecture:

We will implement DeepLabV3+ architecture, which is state-of-the-art for semantic segmentation tasks. We'll use a ResNet-50 backbone pretrained on ImageNet for feature extraction.

Implementation of Framework:

We will implement our network using PyTorch here.

Reference Materials:

We will refer to the following reference materials:

1. Original DeepLabV3+ paper: "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation" (Chen et al., 2018)
2. A2D2 dataset paper: "A2D2: Audi Autonomous Driving Dataset" (Geyer et al., 2020)

Performance Metrics:

We will evaluate our model using the following metrics:

1. **Mean Intersection over Union (mIoU)**: The primary metric for semantic segmentation, measuring overlap between predicted and ground truth segmentation masks.
2. **Per-class IoU**: For selected important classes (e.g., pedestrians, vehicles) to ensure safety-critical objects are well-recognized.
3. **Pixel Accuracy**: The percentage of pixels correctly classified.
4. **Inference Speed (FPS)**: To assess real-time applicability for autonomous driving.

Project Rough Schedule:

- Week 1:
Setup and Data Preparation
- Week 2:
Model Implementation and Training
- Week 3:
Evaluation and Documentation