

GROUP –9 DEEP LEARNING PROJECT PROPOSAL

Multi-Task Learning for Autonomous Driving: Object Detection, Lane Detection, and Traffic Sign Classification

Team Members

Jaya Bharathi Sanjay -11836257

Jaikishan Manivannan -11846539

Manasa Bollinedi -11804584

Sneha Varra -11834552

Problem Statement

Accurate knowledge of several driving scene elements, including lane markings, traffic sign recognition, and vehicle/pedestrian detection, is necessary for autonomous driving. These tasks are typically handled by distinct models, which lengthens the inference time and raises the computational cost. Our objective is to develop an efficient and accurate multi-task learning model that can classify traffic signs, detect objects, and detect lanes.

Methodology

The suggested method extracts common features from input images using a shared convolutional neural network (CNN) backbone. Three task-specific heads are constructed from this common backbone:

Object Detection Head: It Uses a technique modeled after Faster R-CNN or YOLO to identify cars, pedestrians, and other relevant objects.

Lane Detection Head: Accurately detects lanes using a segmentation-based method akin to U-Net or DeepLabv3.

Traffic Sign Classification Head: Uses a lightweight CNN or fully connected layers to identify and categorize traffic signs.

A combined loss function that balances the three tasks will be used to train the model from beginning to end. To enhance generalization, data augmentation techniques such as random flips, rotations, and brightness adjustments will be used. Frameworks like PyTorch or TensorFlow will be used for training in GPU-enabled environments.

Dataset

For every task, we will utilize publicly accessible datasets:

Object detection: bounding box annotations from the KITTI or BDD100K datasets.

Lane Detection: Pixel-by-pixel lane annotations using the TuSimple Lane Detection Dataset.

GTSRB (German Traffic Sign Recognition Benchmark) is used for traffic sign classification.

The CARLA simulator can be used to test the model in a virtual setting and assess how well it performs in actual driving situations.

Baseline

The model will be contrasted with conventional single-task methods.

For object detection, use YOLOv8.

Lane detection with U-Net/DeepLabv3

CNN classifier for recognizing traffic signs

The comparison will concentrate on computational efficiency, inference time, and accuracy.

Evaluation Metrics

Performance will be evaluated by means of:

Mean Average Precision (mAP), recall, and precision in object detection

Lane Detection: Pixel accuracy and intersection over union (IoU)

Classification of Traffic Signs: Precision and F1-score

Frames per second (FPS) and model size as indicators of overall model efficiency

Expected Outcome

The goal is to create a single, unified model capable of performing all three tasks efficiently. We expect to achieve performance comparable to individual models while significantly reducing inference time and resource usage. This approach could contribute to more practical and scalable solutions for autonomous driving systems.