NAVARCH 565 FA 2023 Final Project

Congratulations on finishing the controls half of the Self-Driving Cars course! One more half to go, then we can all start our holidays.

For the final project, we will have a chance to test what we have learned in the perception half of the course with more autonomy. So far we have learned about semantic segmentation, camera calibration, PyTorch neural networks and data loaders, and temporal data. The final project will be on the task of monocular 3D object detection, where the goal is to predict the type, position, size, and orientation of objects from a single RGB image. Figure 1 below shows the type of images you will be making predictions on, similar to your data in Assignment 4. Figure 2 below shows a monocular image with 3D bounding boxes like you will be predicting.

Figure 1: Example image from the final project dataset.

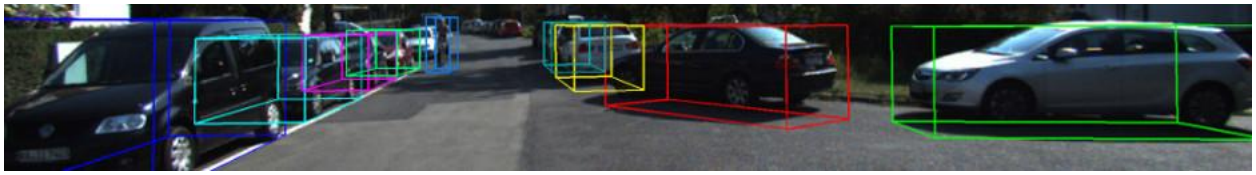

Figure 2: Example of 3D bounding boxes in a monocular image.

Input:
You will be provided a train, validation, and testing split where each split contains one or more driving sequences. Each sequence will contain a file with camera extrinsics, a file with camera intrinsics, a file with object 3D bounding boxes in world and camera coordinates, as well as a folder containing RGB input images and ground truth semantic segmentation labels. Figure 3 demonstrates a ground truth mask for the same driving scene as Figure 1. Each text file contains a header with information about the label present in each column.

All information besides ground truth bounding boxes and semantic segmentation labels will be provided for the test set. It is up to each team to decide how to construct a data loader and train

their network. For example, you may leverage temporal information during training and testing or semantic segmentation ground truth during training.

Figure 3: Example of semantic segmentation label from the same scene as Figure 1. Images can be decoded into integer segmentation values using the same provided code as Assignment 4.

Figure 4: Example of instance segmentation labels from the same scene as Figure 1. Pixel values are equivalent to the TrackID from the bounding box + 1, or 0 if there is no object.

Grading:
Submissions will be evaluated on the hidden test set, with a maximum score of 100% acquired by beating our baseline. Our baseline will be an open source monocular 3D object detection network published in the past 2-3 years, which we train on the dataset by creating a new data loader. You can also use open-source neural networks in the final project, but adaptations and tuning are likely needed to work on this new dataset properly.

Additionally, each group is required to submit one three-page IEEE format paper detailing their approach. Make it clear what is already in the existing implementation and what is newly done by you for the final project if you are using open-source networks. If the groups attempt the extra credit project, the report should also include information on their approach for the extra credit.

Extra Credit:

Perception systems may be very effective in easy driving conditions, but oftentimes we must deal with adverse perceptual challenges. Extra credit for this assignment will be available based on a test driving sequence with adverse weather conditions. An example image is shown in Figure 5.

Figure 5: Example image from extra credit test set.

The extra credit project follows the same format, with the same input and training data provided. The catch: Only the top three teams will receive extra credit, based on a public leaderboard. The top team will receive an extra **30%** on their final project, with 25% for team 2 and 20% for team 3. For this assignment, you may want to read recent papers on adversarial weather and experiment with data augmentation. Good luck!

Link to Data:
Dataset is currently under preparation. We will make a post on Piazza when the dataset is ready for release, and add a link here as well as more instructions. For now, focus on learning from the lectures and homeworks! All skills will be valuable in this assignment.