Course Code	Course Title
COMP 4901V	Large-Scale Deep Perception, Localization, and Planning for Autonomous Vehicles

Course Description

In recent years, autonomous vehicles have become an important field in computer vision and artificial intelligence with the strong support of big visual data and powerful deep learning. This course will deliver a knowledge base for students to understand fundamentals in autonomous vehicles, covering three core modules, i.e., perception, localization, prediction and planning, in large-scale contexts within deep learning. Specifically, (i) for the perception module, we introduce techniques for visual understanding of images and videos captured from RGB/LiDAR cameras, including object recognition, 2D/3D detection, semantic/instance segmentation, motion and flow estimation, and multi-task perception; (ii) for the localization module, we cover topics including camera and geometry models, stereo and monocular depth estimation, 3D scene reconstruction, visual odometry, SLAM, and multi-sensor fusion; (iii) for the prediction and planning module, we introduce pedestrian/vehicle trajectory prediction, vehicle modeling/control, motion and path planning.

List of Topics

1. Introduction
2. Deep Learning Network Architecture and Optimization
   2.1 CNN, RNN, and Transformer
   2.2 Deep Network Optimization
3. Deep Perception
   3.1 Semantic and Instance Segmentation
   3.2 Two-Stage and One-Stage 2D Object Detection
   3.3 Monocular and Point-Cloud 3D Object Detection
   3.5 Optical and Scene Flow Estimation
   3.6 Road, Lane, and Traffic Sign Detection
   3.7. Multi-task Learning and Perception
4. Deep Localization
   4.1 Camera Models
   4.2 Two-view Geometry
   4.3 Stereo Depth Estimation
   4.4 Monocular Depth Estimation
   4.5 3D Scene Reconstruction
   4.6 Visual Odometry and SLAM
   4.7 LiDAR-based Localization
5 Deep Planning
   5.1 Pedestrian and Vehicle Trajectory Prediction
   5.2 Motion Planning and Vehicle Control
   5.3 Path Planning for Autonomous Driving
   5.4 Multi-Sensor Fusion

Textbooks/Reference books

- Deep Learning, Ian Goodfellow and Yoshua Bengio and Aaron Courville
- Computer Vision: A Modern Approach, by David Forsyth and Jean Ponce
- Multiple View Geometry in Computer Vision, by Richard Hartley and Andrew Zisserman.

Grading Scheme

- Two individual project assignments (36%)
- One group project (project proposal, project code & implementation, and project demo & presentation) (24%)
- Final exam (40%)
- Grade by letter

Course Intended Learning Outcomes

The students will build up a knowledge base on large-scale deep-learning-based perception, localization, and planning techniques for autonomous driving, and also be able to transfer the knowledge to design and develop effective vision-based AI solutions to address challenging real problems in our daily life.

Assessment Rubrics

Nil