Adarsh Modh

COMPUTER VISION | MACHINE LEARNING | ROBOTICS

(856) 263-8009 | adarsh.modh@gmail.com | adarshmodh.github.io | adarshmodh | adarshmodh

Education

University of Pennsylvania *Philadelphia, PA*

MASTER OF SCIENCE IN ELECTRICAL AND SYSTEMS ENGINEERING *May 2020*

Coursework - Control Systems, Estimation & Localization in Robotics, Machine Learning, Deep Learning, Machine Perception, Data Mining, Convolutional Neural Networks for Visual Recognition, Reinforcement Learning

S.V. National Institute of Technology *Surat, India*

BACHELOR OF TECHNOLOGY IN ELECTRICAL ENGINEERING *May 2017*

Coursework - Digital Signal Processing, Micro-controllers and Embedded Systems, Mathematics - Linear Algebra, Vector Calculus, Probability

Skills

Programming Languages Python, C++, Embedded-C, MATLAB

Software tools PyTorch, OpenCV, ROS, Gazebo, Autoware, Docker, SLURM, Linux

Embedded System Design Hardware Prototyping for Sensor & Peripheral Interface, Programming 8-bit & 32-bit Micro-controllers (ARM Cortex-M4)

Experience

NEC Laboratories America *San Jose, CA*

RESEARCH ENGINEER, MEDIA ANALYTICS, DR. MANMOHAN CHANDRAKER *December 2020 - Current*

- Leading the development of a robust autonomous vehicle platform, instrumental in creating compelling demonstrations showcasing the lab's research in computer vision and deep learning.
- Currently, I'm focused on the enhancement, deployment, and testing of our software stack, built upon Autoware and ROS in Python and C++.
- This effort primarily involves the replacement of outdated Autoware modules with cutting-edge, in-house implementations, incorporating the lab's pioneering research in areas like Lidar Perception, Prediction, and Planning.
- I have developed a Monocular 3d Object Detection module which works accurately with unconstrained monocular image inputs and using self-supervised techniques this method can be easily adapted to unseen driving, outdoor and indoor scenes. We developed a 2-stage network: in the first stage, a Convolutional Neural Network (Midas) is trained to predict pixel-wise depth from monocular images using supervision from ground-truth LiDAR data. Then, random subsampling is used to convert the dense pixel-wise depth map to a 3D point cloud (pseudo-LiDAR), which mimics a real LiDAR signal. Subsequently, this pseudo-LiDAR signal is fed into a LiDAR-based 3D Object Detection network (CenterPoint) to obtain 3D boxes. This neural network is fine-tuned using the pseudo-LiDAR data and supervised by ground-truth 3D boxes from the KITTI dataset. This approach yielded results comparable to the existing state-of-the-art methods on KITTI data.
- Additionally, I have also worked on another module which does Diverse Lane-aware Trajectory Prediction. My task involved re-implementing the prediction module to provide real-time outputs on a constrained GPU system on our car platform. The original paper implementation was inefficient, sometimes taking approximately 2 seconds to generate predictions per agent. I successfully re-implemented the input processing pipeline by pre-calculating some computationally expensive inputs used for computing the BEVMap tensor, which the network uses to generate outputs. This optimization reduced the time to generate predictions to 0.5 seconds, allowing us to ship this module to our customer in Japan.

GRASP Lab - University of Pennsylvania *Philadelphia, PA*

GRADUATE RESEARCHER - PROJECT: **GEOMETRIC METHODS IN VISION AND LEARNING**, AVIARY - DR. KOSTAS DANIILIDIS *March 2019 - August 2020*

- Played a pivotal role in tackling challenges within multi-view geometry, encompassing tasks such as matching, triangulation, tracking, and re-identification.
- Developed a novel method for establishing multi-view correspondence, thereby assigning 3D consistent IDs to birds inhabiting an aviary. This was implemented by the integration of geometric concepts, specifically Sampson error, and visual models, including a leg-band detector.
- Contributed to the development of a new method 3D Reconstruction for birds (*ECCV 2020*) which involved fitting an articulated 3D mesh model to keypoints and masks within a multiview dataset. This approach leveraged an optimization procedure inspired by SMPLify, a technique initially developed for estimating 3D pose and shape for Humans
- Here, using learning techniques we estimate the 3D pose and shape of birds from a single view. We detect bird instances using a Mask R-CNN pretrained on COCO instance segmentation, which we fine-tune for birds. Given a detection and associated bounding box, we predict body keypoints and a mask. We then predict the parameters of an articulated avian mesh model using pose and shape regression networks, which provides a good initial estimate for further optimization.

Robotics Research Center - IIIT *Hyderabad, India*

RESEARCH ASSISTANT - PROJECT: **PLANNING & CONTROL FOR AUTONOMOUS VEHICLES**, DR. MADHAVA KRISHNA *July 2017 - June 2018*

- Developed an advanced motion planning framework aimed at generating optimized, collision-free trajectories essential for obstacle avoidance and seamless lane merging.
- Implemented a mid-level controller, enhancing the autonomy of the vehicle by efficiently regulating commands from the higher-level planner to the lower-level controller.
- Spearheaded the development of the Drive-by-Wire System which entailed the design and implementation of the lower-level control system for steering, braking, and throttle commands.
- Simulation Videos and Demonstrations on Driverless Car

Academic Projects

Morphology-agnostic Visuomotor Robotic Control

- Developed a novel approach for rapid robotic control within 20 seconds, requiring minimal prior knowledge of the robot's morphology. This method encompasses two core components: Self-Recognition (Learning) and Visual Servoing (Control).
- Self-Recognition: Employed random exploration data, comprising action-observation pairs, to learn a model.
- Visual Servoing: Utilized the acquired model to control the robot, enabling it to perform tasks such as 3D point reaching and trajectory following.

Object Detection

- Successfully implemented YOLO to accurately predict bounding boxes and classify objects within a diverse dataset containing pedestrians, cars, and traffic signals.
- Developed MASK R-CNN from scratch and applied it on the COCO dataset, achieving robust detection of people, animals, and vehicles. This multifaceted task involved the development of a two-stage framework. Initially, I trained a Backbone Region Proposal Network (RPN) to identify candidate object regions. Subsequently, I further refined the model by training the Regressor, Classifier, and Mask Heads, resulting in highly accurate object detection and segmentation.

Estimation and Localization

- Applied Unscented Kalman Filter to estimate the robot's orientation in quaternions, utilizing data from the IMU.
- Executed Landmark-based EKF-SLAM, integrating GPS, LiDAR, and Wheel odometry data from the Victoria Park Dataset to achieve robust simultaneous localization and mapping.

Deep Reinforcement Learning

- Learnt the dynamics for Pendulum-v0 environment in OpenAI gym using Model-Based Learning and then getting the optimal solution with a graph based planner A* over a discretized state-action space.
- Implemented algorithms like DQN and REINFORCE for solving environments like CarRacing-v0 and CartPole-v1.