ME 780
Perception for Autonomous Driving
Spring 2017

Instructor: Professor Steven Waslander

Overview and Objectives

The course will be addressing two concepts:

1. **Deep Learning**: Present a mathematical introduction to deep learning.
2. **Perception For Autonomous Driving**: Review the latest work in perception strategies for autonomous driving.

Schedule

Two 1.5 hours weekly meeting will be held in room TBD. Students in the class will alternate in presenting formal talks on existing research as well as their own findings in simulation using real data sets. Topics for each week of lecture will be agreed upon in the first week of class, and the class is expected to take the standard 12 weeks.

Prerequisites

Students are expected to have good knowledge on linear algebra, probability theory, calculus, numerical computation and traditional machine learning.

Course Outline

The two weekly meeting time will be divided among the two parallel tracks: 1.5 hours for deep learning, and 1.5 hours for perception.

Deep Learning

Students are required to sit for 1.5 hour lectures on deep learning per week.

1. Review of traditional machine learning concepts. (Week 1)
2. Feedforward Neural Networks (Weeks 2,3)
3. Regularization Strategies For Deep Models (Weeks 4,5)

4. Optimization For Training Deep Models (Weeks 6,7)
5. Convolutional Neural Networks (Weeks 8,9)
6. Practical Considerations For Training Deep Models (Week 10)
7. Recurrent Neural Networks (Week 11-12)

Autonomous Driving

Students are required to present two presentations, 1.5 hours each. Each presentation goes in depth through one of the following papers. Furthermore, students will work together to produce a novel 3D object detection algorithm, along with baseline software and an ieee double column 8 page report.

1. Object Proposals Generation and Object Detection
    - Hosang, Jan, et al. "What makes for effective detection proposals?." IEEE transactions on pattern analysis and machine intelligence 38.4 (2016): 814-830.
    - Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. 2015.
    - Chen, Xiaozhi, et al. "3D Object Proposals using Stereo Imagery for Accurate Object Class Detection." arXiv preprint arXiv:1608.07711 (2016).
    - Song, Shuran, and Jianxiong Xiao. "Deep sliding shapes for amodal 3D object detection in RGB-D images." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
    - Chen, Xiaozhi, et al. "Multi-View 3D Object Detection Network for Autonomous Driving." arXiv preprint arXiv:1611.07759 (2016).
    - Chabot, Florian, et al. "Deep MANTA: A Coarse-to-fine Many-Task Network for joint 2D and 3D vehicle analysis from monocular image." arXiv preprint arXiv:1703.07570 (2017).
    - Jimmy Ren, Xiaohao Chen, Jianbo Liu, Wenxiu Sun, Jiahao Pang, Qiong Yan, Yu-Wing Tai, and Li Xu. "Accurate Single Stage Detector Using Recurrent Rolling Convolution." arXiv preprint arXiv:1704.05776 (2017).
2. Semantic and Instance Level Segmentation
    - Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Scene Segmentation." IEEE Transactions on Pattern Analysis and Machine Intelligence (2017).
    - Zhang, Ziyu, Sanja Fidler, and Raquel Urtasun. "Instance-level segmentation for autonomous driving with deep densely connected mrfs." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

- Shelhamer, Evan, Jonathon Long, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." IEEE transactions on pattern analysis and machine intelligence (2016).
- Teichmann, Marvin, et al. "MultiNet: Real-time Joint Semantic Reasoning for Autonomous Driving." arXiv preprint arXiv:1612.07695 (2016).
- Oliveira, Gabriel L., Wolfram Burgard, and Thomas Brox. "Efficient deep models for monocular road segmentation." Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on. IEEE, 2016.
- Wu, Zifeng, Chunhua Shen, and Anton van den Hengel. "Wider or Deeper: Revisiting the ResNet Model for Visual Recognition." arXiv preprint arXiv:1611.10080 (2016).

3. End to End Autonomous Driving:
    - Bojarski, Mariusz, et al. "End to end learning for self-driving cars." arXiv preprint arXiv:1604.07316 (2016).
    - Xu, Huazhe, et al. "End-to-end Learning of Driving Models from Large-scale Video Datasets." arXiv preprint arXiv:1612.01079 (2016).

Grading Policy

Final Project: 70%

The final project is to be done in groups of two. Students are free to use any deep learning package to implement their project. Project topics will be given to students during the first week of the course. Students are to choose one of the topics by the second week of the term.

- Milestone: 5%
- Write-up: 30%
    - Clarity, structure, language, references: 10%
    - Background literature survey, good understanding of the problem: 10%
    - Good insights and discussions of methodology, analysis, results, etc.: 10%
- Technical: 20%
    - Correctness: 8%
    - Depth: 8%
    - Innovation 4%
- Evaluation And Results: 15%
    - Sound evaluation metrics: 5%
    - Thorough analysis and experimentation: 5%

In Class Presentation: 30%

Students are expected to prepare a 30 minute presentation on one of the designated papers to be read in this course. The instructor will be assigning three papers to three students every week.

- Clarity of presentation: 10%
- Analysis of the results: 10%
- Indication of the paper's contribution: 5%
- Keeping to time: 5%