Topics in Machine Perception

CS59200 Seminar Course (MS and PHD)

Course title: Topics in Machine Perception

Course Instructor:
Name: Raymond A. Yeh
Email: rayyeh@purdue.com

Course Description:
This course covers the concepts and techniques for conducting research in the area of machine perception, i.e., how to enable machines to sense the world. The lectures are designed to lead discussions and facilitate student presentations on selected advanced topics in the area. The course aims to develop students' knowledge and analysis capabilities for understanding research publications in machine perception.

Course Schedule (Tentative):

- Motivation / Introduction of machine perception
- Overview of machine perception systems
- Review of machine learning / deep learning
    - Linear Model
    - Convex Optimization
    - Automatic Differentiation
    - Deep Neural Networks
- Machine Perception*
    - Prediction Models (Classification, regression, dense prediction)
    - Deep-Net Building Blocks and Architecture
        - Convolutional Neural Networks Net
        - RNN
        - Transformer
    - Beyond Prediction Models
        - Self-supervised Models
        - Generative Models
    - Multimodal Methods (Vision + X)
    - Applications
        - Audio, Image, Multimodal and 3D

*Lectures followed by students' presentations on a relevant paper.

Required Course Materials:

Recommended Textbooks (Free Available Online):

- Probabilistic Machine Learning by Kevin P. Murphy
- Computer Vision: algorithms and applications 2nd ed. by Szeliski, Richard.

Learning Outcomes:

After this course students should gain the following skills:

- Understand and identify building blocks in a machine perception system; This includes knowledge of:
    - Data representation and pre-processing
    - Training and validating models
    - Machine perception applications
- Effectively read and analyze research papers in the area; This includes knowledge of:
    - Identifying key contributions of a paper
    - Analyzing relation of a paper to prior works
    - Familiarity in the notation and convention used by the community

How to Succeed in this Course:

- Prerequisites:
    - Basic knowledge of linear algebra
    - Basic knowledge of probability
    - Proficiency in Python or can quickly pick it up (e.g., proficiency in MATLAB)
- Responsibilities:
    - Do the readings and homework assignments
    - Participate in lectures and discussions

Grading Criteria:

- Homework: 30%
- Presentation: 30%
- Final Project: 40%

List of Papers (tentative):

- Baevski, Alexei, et al. "data2vec: A General Framework for Self-supervised Learning in Speech, Vision and Language." arXiv preprint arXiv:2202.03555 (2022).
- Radford, Alec, et al. "Learning transferable visual models from natural language supervision." International Conference on Machine Learning. PMLR, 2021.

- Ramesh, Aditya, et al. "Zero-shot text-to-image generation." International Conference on Machine Learning. PMLR, 2021.
- Alexey Dosovitskiy et al. “An image is worth 16x16 words: Transformers for image recognition at scale.” In Proc. ICLR, 2021.
- Chen, Yinbo, Sifei Liu, and Xiaolong Wang. "Learning continuous image representation with local implicit image function." Proc. CVPR. 2021.
- Liu, Ze, et al. "SWIN transformer: Hierarchical vision transformer using shifted windows." Proc. ICCV. 2021.
- Razavi, Ali, Aaron Van den Oord, and Oriol Vinyals. "Generating diverse high-fidelity images with VQ-VAE-2." Proc. NeurIPS, 2019.
- Wang, Yue, et al. "Dynamic graph CNN for learning on point clouds." *TOG,* 2019
- Zamir, Amir R., et al. "Taskonomy: Disentangling task transfer learning." Proc. CVPR. 2018.
- Wu, Zhirong, et al. "Unsupervised feature learning via non-parametric instance discrimination." Proc. CVPR. 2018.
- Zhou, Yin, and Oncel Tuzel. "Voxelnet: End-to-end learning for point cloud based 3d object detection." Proc. CVPR 2018.
- Vaswani, Ashish, et al. "Attention is all you need." NeurIPS, 2017.
- Qi, Charles Ruizhongtai, et al. "Pointnet++: Deep hierarchical feature learning on point sets in a metric space." Proc. NeurIPS, 2017
- Qi, Charles R., et al. "Pointnet: Deep learning on point sets for 3d classification and segmentation." In Proc. CVPR, 2017.
- Van Den Oord, Aaron, and Oriol Vinyals. "Neural discrete representation learning." Proc. NeurIPS, 2017.
- Oord, Aaron van den, et al. "Wavenet: A generative model for raw audio." arXiv preprint arXiv:1609.03499 (2016).
- He, Kaiming, et al. "Deep residual learning for image recognition." Proc. CVPR. 2016.
- Jaderberg, Max, Karen Simonyan, and Andrew Zisserman. "Spatial transformer networks." In Proc. NeurIPS, 2015.
- Hinton, Geoffrey, et al. "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups." IEEE Signal processing magazine, 2012