DEEP LEARNING FOR COMPUTER VISION

PROF. VINEETH N BALASUBRAMANIAN

Department of Computer Science and Engineering

IIT Hyderabad

PRE-REQUISITES :

- Completion of a basic course in Machine Learning
- (Recommended, but not mandatory) Completion of a course in Deep Learning, or exposure to topics in neural networks
- Knowledge of basics in probability, linear algebra, and calculus
- Experience of programming, preferably in Python

INTENDED AUDIENCE : Senior undergraduate students + post-graduate students

INDUSTRIES APPLICABLE TO : All companies that use computer vision for their products/ services (Microsoft, Google, Facebook, Apple, TCS, Cognizant, L&T, etc)

COURSE OUTLINE :

The automatic analysis and understanding of images and videos, a field called Computer Vision, occupies significant importance in applications including security, healthcare, entertainment, mobility, etc. The recent success of deep learning methods has revolutionized the field of computer vision, making new developments increasingly closer to deployment that benefits end users. This course will introduce the students to traditional computer vision topics, before presenting deep learning methods for computer vision. The course will cover basics as well as recent advancements in these areas, which will help the student learn the basics as well as become proficient in applying these methods to real-world applications. The course assumes that the student has already completed a full course in machine learning, and some introduction to deep learning preferably, and will build on these topics focusing on computer vision.

ABOUT INSTRUCTOR :

Vineeth N Balasubramanian is an Associate Professor in the Department of Computer Science and Engineering at the Indian Institute of Technology, Hyderabad (IIT-H). He was also the Founding Head of the Department of Artificial Intelligence at IIT-H from 2019-22, and a Fulbright-Nehru Visiting Faculty at Carnegie Mellon University in 2022-23. His research interests include deep learning, machine learning, and computer vision. His research has resulted in over 160 peer-reviewed publications at various international venues, including top-tier venues such as ICML, CVPR, NeurIPS, ICCV, KDD, AAAI, and IEEE TPAMI, with Best Paper Awards at recent venues such as CODS-COMAD 2022, CVPR 2021 Workshop on Causality in Vision, etc. He served as a General Chair for ACML 2022, and serves as a Senior PC/Area Chair regularly for conferences such as CVPR, ICCV, AAAI, IJCAI and ECCV. He is a recipient of the Google Research Scholar Award (2021), NASSCOM AI Gamechanger Award (2022, both Winner and Runner-up), Teaching Excellence Award at IIT-H (2017 and 2021), Research Excellence Award at IIT-H (2022), among others. For more details, please see https://people.iith.ac.in/vineethnb/.

COURSE PLAN:

Week 1: Introduction and Overview:

- Course Overview and Motivation; History of Computer Vision; Image Representation; Linear Filtering, Correlation, Convolution; Image in Frequency Domain
- (Optional) Image Formation; Image Sampling

Week 2: Visual Features and Representations:

- Edge Detection; From Edges to Blobs and Corners; Scale Space, Image Pyramids and Filter Bank; SIFT and Variants; Other Feature Spaces
- (Optional) Image Segmentation, Human Visual System

Week 3: Visual Matching:

- Feature Matching; From Points to Images: Bag-of-Words and VLAD Representations; Image Descriptor Matching; From Traditional Vision to Deep Learning
- (Optional) Hough Transform; Pyramid Matching

Week 4: Deep Learning Review:

- Neural Networks: A Review; Feedforward Neural Networks and Backpropagation; Gradient Descent and Variants; Regularization in Neural Networks; Improving Training of Neural Networks

Week 5: Convolutional Neural Networks (CNNs):

- Convolutional Neural Networks: An Introduction; Backpropagation in CNNs; Evolution of CNN Architectures for Image Classification; Recent CNN Architectures; Finetuning in CNNs

Week 6: Visualization and Understanding CNNs:

- Explaining CNNs: Visualization Methods; Early Methods (Visualization of Kernels; Backprop-to-image/Deconvolution Methods); Class Attribution Map Methods (CAM,Grad-CAM, Grad-CAM++, etc); Going Beyond Explaining CNNs
- (Optional) Explaining CNNs: Recent Methods

Week 7: CNNs for Recognition, Verification, Detection, Segmentation:

- CNNs for Object Detection; CNNs for Segmentation; CNNs for Human Understanding: Faces
- (Optional) CNNs for Human Understanding: Human Pose and Crowd; CNNs for Other Image Tasks

Week 8: Recurrent Neural Networks (RNNs):

- Recurrent Neural Networks: Introduction; Backpropagation in RNNs; LSTMs and GRUs; Video Understanding using CNNs and RNNs

Week 9: Attention Models:

- Attention in Vision Models: An Introduction; Vision and Language: Image Captioning; Self-Attention and Transformers
- (Optional) Beyond Captioning: Visual QA, Visual Dialog; Other Attention Models

Week 10: Deep Generative Models:

- Deep Generative Models: An Introduction; Generative Adversarial Networks; Variational Autoencoders; Combining VAEs and GANs
- (Optional) Beyond VAEs and GANs: Other Deep Generative Models

Week 11: Variants and Applications of Generative Models in Vision:

- GAN Improvements; Deep Generative Models across Multiple Domains; Deep Generative Models: Image Application
- (Optional) VAEs and Disentanglement; Deep Generative Models: Video Applications

Week 12:Recent Trends:

- Few-shot and Zero-shot Learning; Self-Supervised Learning; Adversarial Robustness; Course Conclusion
- (Optional) Pruning and Model Compression; Neural Architecture Search