



# DEEP LEARNING FOR COMPUTER VISION

**PROF. VINEETH N BALASUBRAMANIAN**

Department of Computer Science and Engineering  
IIT Hyderabad

**TYPE OF COURSE** : New | Elective | UG/PG

**COURSE DURATION** : 12 weeks (20 Jul' 20 - 9 Oct' 20)

**EXAM DATE** : 18 Oct 2020

**PRE-REQUISITES** : Completion of a basic course in Machine Learning Completion of a basic course in Deep Learning Knowledge of basics in probability, linear algebra, and calculus

**INTENDED AUDIENCE** : Any Interested Learners

**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 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.

## **ABOUT INSTRUCTOR :**

Prof. Vineeth N Balasubramanian is an Associate Professor in the Department of Computer Science and Engineering at the Indian Institute of Technology, Hyderabad.

## **COURSE PLAN :**

- Week 1:** Introduction and Overview: Course Overview and Motivation; Introduction to Image Formation, Capture and Representation; Linear Filtering, Correlation, Convolution
- Week 2:** Visual Features and Representations: Edge, Blobs, Corner Detection; Scale Space and Scale Selection; SIFT, SURF; HoG, LBP, etc.
- Week 3:** Visual Matching: Bag-of-words, VLAD; RANSAC, Hough transform; Pyramid Matching; Optical Flow
- Week 4:** Deep Learning Review: Review of Deep Learning, Multi-layer Perceptrons, Backpropagation
- Week 5:** Convolutional Neural Networks (CNNs): Introduction to CNNs; Evolution of CNN Architectures: AlexNet, ZFNet, VGG, InceptionNets, ResNets, DenseNets
- Week 6:** Visualization and Understanding CNNs: Visualization of Kernels; Backprop-to-image/Deconvolution Methods; Deep Dream, Hallucination, Neural Style Transfer; CAM, Grad-CAM, Grad-CAM++; Recent Methods (IG, Segment-IG, SmoothGrad)
- Week 7:** CNNs for Recognition, Verification, Detection, Segmentation: CNNs for Recognition and Verification (Siamese Networks, Triplet Loss, Contrastive Loss, Ranking Loss); CNNs for Detection: Background of Object Detection, R-CNN, Fast R-CNN, Faster R-CNN, YOLO, SSD, RetinaNet; CNNs for Segmentation: FCN, SegNet, U-Net, Mask-RCNN
- Week 8:** Recurrent Neural Networks (RNNs): Review of RNNs; CNN + RNN Models for Video Understanding: Spatio-temporal Models, Action/Activity Recognition
- Week 9:** Attention Models: Introduction to Attention Models in Vision; Vision and Language: Image Captioning, Visual QA, Visual Dialog; Spatial Transformers; Transformer Networks
- Week 10:** Deep Generative Models: Review of (Popular) Deep Generative Models: GANs, VAEs; Other Generative Models: PixelRNNs, NADE, Normalizing Flows, etc
- Week 11:** Variants and Applications of Generative Models in Vision: Applications: Image Editing, Inpainting, Superresolution, 3D Object Generation, Security; Variants: CycleGANs, Progressive GANs, StackGANs, Pix2Pix, etc
- Week 12:** Recent Trends: Zero-shot, One-shot, Few-shot Learning; Self-supervised Learning; Reinforcement Learning in Vision; Other Recent Topics and Applications