CS101: Fundamentals of Computer Vision Programming

Jimut Bahan Pal

Semester - II, 2024

E-mail: pal.jimut@iitb.ac.in
Website: jimut123.github.io/courses/FCVP/sem1_2024.html

Course Description

The integration of computer vision algorithms into everyday activities is on the rise. This course begins with foundational Python3 programming, progressing to the practical application of programming concepts through the widely-used OpenCV library, known for classical computer vision approaches. Participants will delve into the PyTorch framework, mastering the design of models tailored for classification and image segmentation—key algorithms in the industry and research. Additionally, there is potential for hands-on experience in developing a real-world deployable app, showcasing the practical application of acquired skills.

This introductory course combines theoretical concepts with hands-on tutorials, catering to individuals keen on delving into computer vision. The practical, tutorial-driven approach is tailored for those eager to apply classical and prevalent deep learning algorithms to real-world scenarios, enhancing their understanding and skill application in the field.

Required Materials

- Course materials, e.g., slides, jupyter notebook would be made available via Google Classrooms.
- Pointers to additional resources week-wise will be made via course website or via Google classrooms.

Prerequisites

No formal prerequisites; only a willingness to learn and explore.

Course Objectives

Upon completion, students will demonstrate proficiency in coding computer vision algorithms using Python3 and OpenCV. Additionally, they will possess a foundational understanding of PyTorch and preliminary deep learning algorithms relevant to both research and industry applications.

Grading Policy

(Tentative, total - 110%) Relative grading will be used.

- **20%**: Assignments (about 2-3).
- **30%**: Course Project.
- **15%**: Recitation.
- **10%**: Challenge Programming Assignment.
- **15%**: Mid-semester examination.
- **20%**: End-semester examination.

Tentative Topics

Part - 1: Python3 Refresher

Lecture 1: Python3 Programming Refresher.

Part - 2: OpenCV Details

Lecture 2: Installation, GUI features, Basic Operations, Colorspaces, Geometric Transformations, Image Thresholding, Image Smoothing.

Lecture - 3: Morphological Operations, Image Gradients, Canny Edge, Detection, Image Pyramids.

Lecture - 4: Conoutours, Histogram, Fourier Transform.

Lecture - 5: Template matching, Hough transform - Line, Hough transform - Circle, Watershed, Grab-Cut.

Lecture - 6: Feature detection, Harris Corner, Shi-Tomasi Detector, SIFT, SURF.

Lecture - 7: FAST, BRIEF, Feature Matching, Homography.

Lecture - 8: Video (Optical Flow, Background subtraction methods, Meanshift and Camshift).

Lecture - 9: HaarCascade, Camera Calibration, Pose Estimation, Epipolar Geometry, Depth Map from stereo image.

Lecture 10: K-Means Clustering, K-Nearest Neighbor, Image Denoising, Image Inpainting, High Dynamic Range (HDR).

Part - 3: Pytorch Fundamentals

Lecture 11: Introduction to Pytorch and Tensors, Feed Forward Neural Networks, Back-propagation, Gradient descent, Classifier example, Activation functions.

Lecture 12: Optimizers, Convolutional Neural Networks, Convolution operation, Maxpool.

Lecture 13: Basic Methodology - Train - Validation - Test Pipeline, Design a Classifier using Pytorch, Autoencoders - theory and example.

Lecture 14: Segmentation - Single class and Multi-Class, Theory: U-Net, Attention U-Net, ResU-Net, DeepLabV3+.

Lecture 15: Application - Design a Segmentation Network using Pytorch.

Lecture 16: VAEs - Theory.

Lecture 17: Application - Design a VAE using Pytorch.

Lecture 18: GANs - Theory.

Lecture 19: Application - Design a GAN using Pytorch.

Lecture 20: (Tentative) Guest Lecture: Transformer & Vision Transformer Theory and Tutorial, Seshadri Mazumder, CVIT, IIIT Hyderabad.

Lecture 21: (Tentative) Guest Lecture: Stable Diffusion Theory and Tutorial, Seshadri Mazumder, CVIT, IIIT Hyderabad.