# DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

# KATHMANDU UNIVERSITY

**Subject: Computer Vision [Elective]** **Course Code: COMP 493**

**Credit: 3** **Full Marks: 100**

## Course Description:

This course is intended to provide students with an entry point into the fundamental ideas and significant applications of computer vision. Several key ideas in computer vision will be introduced during the course. We will introduce students to several real-world applications that are crucial to our everyday lives. More significantly, we will mentor students through a number of thoughtful projects so they can put a few intriguing and cutting-edge computer vision algorithms into practice.

## Objectives:

Upon completion of this course, students should be able to:

- Recognize and describe both the theoretical and practical aspects of computing with images. Connect issues from Computer Vision to Human Vision
- Describe the foundation of image formation and image analysis. Understand the basics of 2D and 3D Computer Vision
- Become familiar with the major technical approaches involved in computer vision. Describe various methods used for basic and advance image pre-processing
- Get an exposure to advanced concepts leading to object categorization and segmentation in images
- Build computer vision applications

## Pre-requisites:

It is not assumed that the students have any prior computer vision experience, however prior exposure to visual computing or signal processing will be beneficial. For this course, following skills is needed: Data structures, Programming, Math (Linear algebra, vector calculus, and probability)

## Evaluation:

Internal – 50%

Final – 50%

## Course Contents:

**Unit 1: Introduction** **[4 hours]**

**1.1** Human Vision
**1.2** Color Spaces and Transformations

**Unit 2: Image Processing** **[6 hours]**

**2.1** Image coordinates and Resizing

**2.2** Filters and convolutions

**2.2.1**Point operators

**2.2.2** Linear filtering

**2.2.3**Non-linear filtering

**2.2.4** Fourier transforms

**2.2.5** Pyramids and wavelets

**2.2.6** Geometric transformations

**Unit 3: Interpolation and Optimization** **[5 hours]**

**3.1** Scattered data interpolation
**3.2** Variational methods and regularization
**3.3** Markov random fields

**Unit 4: Learning Paradigm for Computer Vision** **[8 hours]**

**4.1** Machine Learning
4.2. Neural Networks
4.3. Convolutional Neural Networks and Other Network Architectures
4.4. Transfer Learning
4.5 Autoencoder

**Unit 5: Object Detection and Semantic Segmentation** **[5 hours]**

**5.1** Object Detection

**5.2** Semantic Segmentation

**5.3** Deep Learning Frameworks for Object detection and Segmentation

**5.4** Panoptic Segmentation and Video

**Unit 6: Edges, features and Processing** **[6 hours]**

6.1 Edges and its features

6.2 Edge Matching and RANSAC phenomenon

6.3 Edge Processing algorithms

6.3.1 Histogram of Oriented Gradients (HoG)

6.3.2 Scale Invariant Feature Transformations (SIFT)

6.3.3 Optical Flow

**Unit 7: 3-D vision** **[5 hours]**

**7.1** Introduction

**7.2** Depth Perception

**7.3** Stereo Vision

**Unit 8: Advanced topics in Computer Vision** **[6 hours]**

**8.1** Simultaneous Localization and Mapping (SLAM/SFM)

**8.2** 3D shape/appearance model

**8.3** Computational photography

**8.4** Generative Adversarial Networks (GANs)

**8.5** Vision Transformers

**Text Book:**

Computer Vision: Algorithms and Applications, 2nd ed., Richard Szeliski, 2022, Springer. ISBN-13: 978-3030343712

**Reference Books:**

Computer Vision: A Modern Approach 2nd Edition, David Forsyth and Jean Ponce, ISBN-13: 978-0136085928

Multiple View Geometry in Computer Vision, Second Edition, Richard Hartley and Andrew Zisserman, Cambridge University Press, March 2004.