Fusemachines Research and Training Center

Syllabus

Computer Vision

Microdegree in Artificial Intelligence Program

Version	Significant Changes (Marked with a Symbol)	Modified by	Modification Date
1.0	First Version of the course	Pradip, Rojesh	

Contents

Introduction

All the syllabuses are reviewed regularly so that it will be able to reflect the latest thinking and current best practices employed in industries and take into account different national and international contexts in which these courses will be taught.

Syllabus Aims

The Syllabus aims to:

- Provide a worthwhile learning experience for all learners and enable them to acquire sufficient knowledge and skills to get started in the domain of AI
- Facilitate and Standardise Course Content Development and Delivery

Introduction to the Course

This is the second part of the Microdegree course covering different topics in Deep Learning.

Tentative Teaching Guide

12 weeks course maximum of 4 hours per week.

Day	Units Covered	Teaching Guide (Instructions for Instructors)	Self-Study Hrs
1	1.1, and 1.2		
2	2.1 and 2.2		
3	3.2.1, 3.2.2 and 3.2.3		
4	3.2.4, 3.3.5 and 3.3.1		
5	4.2 and 4.3		
6	4.4.1, 4.4.2.A(e) and 4.4.2.B(a)		
7	4.5 and 4.6		
8	5.2.1.H and 5.3		
9	6.2 and 6.3		
10	6.4 and 6.5		
11	6.6 and 6.7		

12	Mid-Examination		
13	7.1 and 7.2.1-7.2.3		
14	7.2.4, 7.2.5 and 7.3		
15	8.1 and 8.2		
16	8.3 and 8.3.2.A.d		
17	9.1 and 9.2		
18	9.3 and 9.4		
19	10.1 and 10.2		
20	10.3, 10.4 and 10.5		
21	11.1 and 11.2		
22	11.3		
23	Final Project Presentation		
24	Examination		

Pre-requisites

Students must have completed the Machine Learning and Deep Learning course.

Distinct Features Used in the Syllabus

Bold Outcomes refers to **Must Have** learning outcomes,

Normal Text refers to Should Have learning outcomes,

Italic Outcomes refer to *Good to Have or Could Have learning outcomes*.

Red Outcomes refers to Won't have Outcomes

Colour based on Categories

Bold Dark Blue Teach **Fundamental Concepts**

Bold Dark Green Covers **State-of-the-Art Research Topics**

Bold Dark Orange Focuses on **Application in different fields** the topic is being used.

Indicators:

Indicator	**Indicates**
LO#.#.# (eg: Re LO2.2.2.A)	Learning Objective 'A' of chapter 2 of Module 2, unit 2
Re	Revisit or Relate to

(D)	Demo/ Walkthrough
AO# (eg: AO3)	Focusing on Assessment Objective 3, i.e to apply
(W /P /G /I)	Possible Activities Whole class / Group work/ Pair Activities / Individual Activities
(H)	Homework to be covered in assignments or Projects
(F)	Formative assessment

Syllabus Outline:

Computer Vision

Recommended Books

1.	Digital Image Processing 4th ed, Gonzalez and Woods	
2.	Computer Vision, Algorithms and Applications, Szeliski	
3.	Forsyth & Ponce, Computer Vision: A Modern Approach	
4.	Multiple View Geometry in Computer Vision, 2nd ed, Hartley & Zisserman	

Assessment

The Assessment will be based on the cognitive domain of Bloom's Taxonomy to classify learning objectives into different levels of complexity and specificity, viz. Remember, Understand, Apply, Analyze, Evaluate, Create

Assessment Objectives	Categories	Objective	Action Words
AO1	**Remember**	Recall Facts and basic concepts	Define, list, memorize, repeat, state, recognize
AO2	**Understand**	Explain Ideas or Concepts	Classify, describe, discuss, explain, identify, locate, recognize, report, select, translate, interpret, exemplify,
AO3	**Apply**	Use Information in a new situation, mathematical modeling	Execute, implement, solve, use, demonstrate, interpret, operate, schedule, sketch
AO4	**Analyze**	Draw connections among ideas	Differentiate, organize, relate, compare, contrast, distinguish, examine, experiment, question, test
AO5	**Evaluate**	Justify a stand or decision	Check, Appraise, argue, defend, judge, select, support, value, critique, weigh
AO6	**Create**	Produce new or original Work	Design, assemble, construct, conjecture, develop, generate, plan, produce, formulate, investigate

Source: https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/

Weight Distribution on

Components	AO1	AO2	AO3	AO4	AO5	AO6	Weights
Quizzes	20%	20%	20%	20%	10%	-	
Programming Assignment	10%	20%	20%	20%	10%	10%	
Projects	-	-	10%	20%	20%	50%	
Classroom Assessment							

Course Contents

1. Introduction to Computer Vision (CV)

1.1. Introduction to the Course

Students should be able to :

1. Introduction to the Course	A. Describe the overall goals of Computer vision and image understanding B. Justify that computer vision is multidisciplinary exemplified by different algorithms and applications of computer vision with popular vision systems in use C. Differentiate between Image processing, Computer vision and Computer Graphics D. Recall what course covers and what it does not cover	
2. Course Logistics	A. Recall and follow the assessment and evaluation criteria B. Recall and follow the honor code and violation issues	
1. Brief History of Computer Vision	A. List different application areas of computer vision a. Autonomous driving b. Medical Imaging c. Drones d. Manufacturing e. Biometrics (face, fingerprint etc.) f. Surveillance B. Explain why computer vision is difficult, in terms of being inverse problem and explaining the different challenges such as a. Scale b. Viewpoint variation c. Non-Rigid Deformation d. Intra-object/class variations e. Illumination f. Color and illusions g. Occlusion h. Background Clutter and Camouflage i. Motion j. Local ambiguity	

1.2. Image formation

Students should be able to :

1. Imaging	A. Describe Image Formation using Lambertian Model, as a function of shape, illumination, and reflectance B. Describe pinhole perspective camera C. Describe camera with lens for gathering light, depth of field (or depth of focus), and challenges (blur vs. noisy images) D. Describe Projection Models (3D to 2D) a. Perspective Projection (Pinhole Model) and similarity triangle relation b. Orthographic Projection (UAV Images)	

2. Imaging Geometry	A. Use Projective Transformation with homogeneous coordinates to perform Translation, Scaling, Rotation and their inverses B. Describe how Intrinsic and Extrinsic Parameters affect imaging (Details on how they can be recovered should be covered later.) C. Use Perspective Transformation to relate homogeneous world coordinate to homogeneous image coordinate. And the inverses. D. Rodrigue’s Formula E. Quaternions	
3. Digital Camera	A. Describe different blocks and concepts in digital camera: a. Optics (Lens) b. Aperture c. Shutter d. Sensor e. Gain- ISO f. A/D converter g. Demosaic h. f-number i. Exposure/Integration time j. Vignetting B. Describe different Image Sensors a. CCD vs CMOS sensors and camera setup b. Quantum efficiency C. Camera optics a. Describe aberrations: lens distortion and chromatic aberration. b. Motion-blur and noise and relation to integration time c. Color Filters: R, G and B d. Understand Gray and RGB camera e. Understand Bayer filter mosaic used by RGB camera D. Discuss about Image Sampling and Quantization and represent Pixel as a function E. Describe different types of noise and sources and how they can be modelled: a. Shot noise due to photon counting process with Gaussian (with central limit theorem) b. Salt and pepper noise c. Random noise due to ISO d. Fixed pattern noise e. Quantization noise	
4. Color	A. Describe human color perception with sensitivity to light wavelengths and color receptors B. Describes physics of color C. Color Image theory a. Tristimulus theory b. Colourimetric equation D. Luminosity function	

	E. Perception based colour models: CIE RGB and CIE XYZ F. Additive and subtractive colour models: RGB and CMY G. Luminance and chrominance colour models: YUV, YIQ, YCbCr H. Additive perceptual colour models: HSV and HLS I. Implement Image color transformation RGB to Gray, YUV, HSV using OpenCV	
~~Theories of Vision~~	~~A. Explain mechanisms underlying visual object recognition in the brain~~ ~~a. Marr's theory of vision~~ ~~b. Ventral Visual Stream study in Monkey~~ ~~c. Core object recognition~~ ~~B. Compare Primate Vision, Human Vision and Machine Vision~~ ~~C. Understand important functions and limitations of human vision~~	~~Mechanisms Underlying Visual Object Recognition: Humans vs. Neurons vs. Machines~~ ~~Figure1 | Seeking Categories in the Brain~~
5. Basics for Multi-view geometry	A. Geometric primitives a. 2D point b. 3D point c. Describe a point in world coordinate and image coordinate B. Describe transformations 2D and 3D a. Translation b. Rotation c. Scaling C. 3D to 2D projections a. Perspective projection b. Camera matrix i. Intrinsic parameters ii. Extrinsic parameters D. Multi-view geometry a. Two view geometry i. Epipolar geometry 1. Epipoles 2. Epipolar line b. Map a 3D point from one camera to other camera c. Fundamental Matrix d. Homography E. Robust model estimation a. RANSAC b. LMEDS	

2. Image Processing

Introduction to the Module

Students should be able to :

1. Introduction to the	A. Understand image data, color formats and conversion between the formats	

Module	B. Understand and use image filtering as a function of the local neighborhood for various applications C. Understand image processing in spatial and frequency domains D. Understand multi-resolution image E. Understand image assessment techniques F. Understand the flow of the contents in the module (overview) and their learning outcomes	

2.1. Image Operators

1. Point Operators	A. Plot histogram of image and locate/estimate mean and standard deviation
	B. Convert image to other color spaces and view each component/channel
	  a. YUV
	  b. Lab
	  c. HSV
	C. Define and Implement simple point/pixel operators:
	  a. Brightness and contrast adjustment with scaling and by adding bias
	  b. Brightness and contrast adjustment with gamma correction
	  c. Brightness and contrast adjustment with histogram equalization or tone mapping
	  d. Color balancing
	D. Define Blend Operator and blend two images
	E. Implement Image matting and compositing as blue/green
		Szeliski 3.1
		https://docs.opencv.org/3.4/d3/dc1/tutorial_basic_linear_transform.html

2. Convolutions and cross-correlations	A. Introduce convolution and discuss properties:
	  a. Commutative
	  b. Associative
	  c. Distributive
	  d. Scaling
	  e. Shift invariance
	B. Introduce convolution two dimension
	C. Padding techniques
	  a. Zero
	  b. Constant
	  c. Clamp
	  d. Wrap	

3. Spatial domain processing	A. Introduce spatial domain processing and some examples of spatial frequencies/details in images
	B. Linear filtering
	  a. Explain 2D Gaussian filter and plot the filter response for various mean and sigma values
	  b. Show in spatial domain 2D convolution with 2D Gaussian filter
	  c. Introduce separable filters for 2D Gaussian filter and discuss saving on computational cost
	  d. Explain Gabor filter
	  e. Use OpenCV API to apply Box, Canny, Sobel	

	filters and visual results C. Image derivatives (High and Bandpass filtering) a. Revisit idea of forward, backward and central difference techniques to compute derivatives b. Image derivatives of first and second orders with finite difference technique c. Show how noise is amplified with finite difference technique d. Apply Gaussian filtering for noise reduction followed by finite difference e. Image derivative with the derivative of the gaussian filter (combine derivative and gaussian filter using associative rule) f. Plot filter responses of derivative kernels and derivative of Gausisian (for first and second orders)	
4. Frequency domain processing	A. Introduce frequency domain processing with fourier transform B. Computer convolution in fourier domain and compare execution time with convolution in spatial domain C. Wiener filtering D. Properties of fourier transform a. Shift invariance b. Rotation c. Frequency scaling d. Importance of phase E. Other transforms a. Discrete cosine transforms b. Discrete hartley transform F. Wavelets a. Gabor b. Haar G. Use OpenCV API to perform Wiener filtering	
5. Scale-space or multiresolution image processing	A. Describe multi-resolution representation of image with image pyramid: coarse, medium and fine details B. Visual images at different scales and derivatives C. Describe some applications a. Feature tracking in image pyramid b. Multi-band blending of two images	
6. Integral Image	A. Introduce the concept of integral image B. Feature computation using box filter in face detection originally used by Viola C. Discuss approximation and integral image to accelerate convolution: a. Approximation of Gaussian derivative filters in Hessian as in SURF b. Approximation of first order Gaussian derivative filter in LOCOCO	SURF: http://people.ee.ethz.ch/~surf/eccv06.pdf LOCOCO: https://ieeexplore.ieee.org/document/5727936

2.2. Image restoration

1. De-noising	A. Introduce the concept of image de-noising a. Describe models for various sources of noise covered in Sec. 1.1 b. Introduce noise into some example images e.g. additive gaussian noise, salt and pepper noise, etc. B. Use OpenCV API to show image noise removal for noise given in Sec. 1.1 a. Gaussian smoothing b. Box filter smoothing c. Median filtering d. Bilateral filtering e. Non-local means	
2. Sharpening	A. Introduce the concept of sharpening/deblurring a. Describe the role of lens for blurring and model the system with Gaussian b. Describe the role of integration time for motion blurring B. Describe models for a. out-of-focus blur (Gaussian) b. motion blur (box filter) c. Introduce corresponding blur types into some example images and visualize C. Use OpenCV API to show image sharpening a. Out-of-focus deblurring using unsharp mask	
3. Image upscaling and downscaling	A. Introduce image downscaling a. Describe aliasing problem b. Describe smoothing and downsampling B. Introduce image upscaling with interpolation techniques a. Linear interpolation b. Bilinear c. Bicubic C. Use OpenCV API to show downscaling and upscaling operations by x2 and x4 times	
4. Image quality assessment metrics	A. Describe image assessment techniques a. PSNR b. SSIM B. Use OpenCV API to demonstrate the use of SSIM with ground truth and recovered images C. Deep learning based perceptual metrics a. LPIPS	https://arxiv.org/pdf/1801.03924.pdf https://github.com/richzhang/PerceptualSimilarity
Assignment	A. Use non-local means to de-noise the image.	

2. Feature Detection and Matching

2.1. Introduction to the Module

Students should be able to :

1. Introduction to the Module	2. Understand different local operators to detect local features such as edge, corner and blobs. 3. Understand the importance of edge and feature detection as one of the most important lower-level computer vision for various applications	

2.2. Local Features

Students should be able to :

1. Edge Detectors	A. Describe various edge detectors a. Prewitt edge b. Sobel c. Canny d. Laplacian B. Use OpenCV API to apply these detectors and visualize the results	Szeliski 3.1
2. Feature Detection (e.g. blob and corners, ..)	A. Describe salient features: corner, blob. B. Describe feature invariance for: translation, planar rotation, scaling, affine, perspective, noise, blur C. Describe Harris corner detector a. Explain cornerness measure with eigenvalues related with trace and determinant b. Use OpenCV API to detector corner in image and visualize results D. Feature detectors a. Lapalacian and Difference-of-Gaussian b. Hessian c. Gabor (e.g. SIFER) E. Scale-invariant detectors a. Describe SIFT feature detector i. Describe scale-scape pyramid with Gaussian smoothing ii. Describe Difference of Gaussian for feature detection and explain how it approximates Laplacian of Gaussian and relation to invariance b. Describe SURF feature detector i. Describe approximation of second order gaussian derivatives with box kernel ii. Explain Hessian feature detector c. Use OpenCV API to detect features in the image and also compare execution times F. Intensity based detectors a. Describe MSER detector b. Describe FAST detector c. Describe ORB G. Describe feature invariance for: translation, planar rotation, scaling, affine, perspective, noise, blur H. Describe feature repeatability to measure the quality of feature detectors for invariance	SIFER: Scale-Invariant Feature Detector with Error Resilience

3. Feature Descriptor	A. Filtering based descriptors a. Describe SIFT feature descriptor b. Describe SURF feature descriptor c. Describe PCA-SIFT feature for dimensionality reduction B. Intensity based descriptors a. Describe ORB descriptor and BRIEF b. Describe FREAK c. Describe DAISY descriptor C. Describe feature discriminability to measure the quality of feature descriptors	
4. Feature Matching	A. Describe nearest neighbor distance ratio (NNDR) feature matching B. Describe locality sensitive hashing technique for feature matching C. Describe k-d tree and FLANN for for efficient feature matching D. Use OpenCV API to demonstrate feature matching using AADR and FLANN.	
5. Applications	A. Image Registration or Feature based alignment a. Describe image registration application b. Briefly describe Homogrpahy and RANSAC c. Use OpenCV API using SIFT for image registration i. Detect feature and show results on both the images ii. Match image features and show correspondences on the images iii. Use RANSAC to estimate Homography model iv. Generate registered images B. Describe panorama creation using OpenCV API a. Follow all the steps in image registration b. Describe some blending techniques c. Use blending to create panorama C. Describe basics of Camera Calibration D. Describe basics of Pose Estimation	
Assignment for the unit	A. User feature based technique for video stabilization for a hand held camera.	

2.3. Morphological operations

Students should be able to :

1. Morphological Processing	A. Morphological operations B. Grey level morphology C. Grey level dilation and erosion D. Minkowski operators	
Assignment	A. Use OpenCV API	

	a. Convert image to grey level b. Demo erosion for detecting contours	

2.4. Module Summary

Students should be able to :

Project Ideas for the Module	A. Robot navigation (Visual SLAM) a. Robots need to localize themselves while exploring navigation. b. Use local features to estimate pose of the robot to support navigation and path finding c. Estimate pose B. Camera calibration (3D vision) a. Calibrating the camera (intrinsic and extrinsic) is important in 3D vision b. Use one of the feature detectors to calibrate the camera	

3. Image Classification, Recognition and Matching

3.1. Introduction to the Module

Students should be able to :

Introduction	A. Introduce image classification task as finding different categories of objects (e.g. car, mug, dog, cat, ...) B. Introduce image detection as finding objects and their location (where the objects/faces are in image) C. Introduce recognition task as identifying the face (e.g. whose face e.g. Geoffrey Hinton) D. Introduce image matching as a task whether the given two images are similar or not	
Applications	A. Autonomous driving B. Medical C. Cybersecurity D. Drones E. Robots F. Face matching in smartphone	

3.2. Traditional Image Classification, Detection, Matching and Recognition

Students should be able to :

1. Introduction	A. Introduce building blocks of traditional object detection	

	a. Feature engineering b. Machine learning algorithm for classification B. Motivation of feature engineering for invariance to: noise, blur, scale, light, contrast, etc. C. Challenges and limitations of hand-engineered features and motivation for deep learning	
2. Classification, detection, recognition and matching algorithms	A. Image classification a. Image classification using bag of words (e.g. SIFT features) B. Detection algorithms a. Viola's face detection algorithm describing feature extractor and Adaboost for classification b. Pedestrian Detection using HOG features and SVM C. Recognition a. Face recognition with Eigenfaces D. Matching a. Image retrieval/matching in database using bag of words (e.g. SIFT features)	
3. Assignments	A. Use SIFT features to create bag of words. Classify some object categories by using the method.	

3.3. Dataset Preparation Tools and Public Datasets for Computer Vision

Students should be able to :

1. Data preparation tools	A. Recap a need for data preparation, annotation and labeling for computer vision, a. Classification needs: picture + label b. Detection: bounding box + label c. Segmentation: polygon + label d. 3D cuboids (eg: SDC) e. Key-Point and Landmark f. Lines and Splines (eg: lanes detection) B. Discuss on different type of annotation and annotation and data preparation tools and format for computer vision a. Visual Object Tagging Tool (VoTT) (https://github.com/microsoft/VoTT) b. VGG Image Annotator (VIA) (Visual Geometry Group) c. Computer Vision Annotation Tool (CVAT) (https://github.com/opencv/cvat) d. Labelbox e. LabelMe. The Open annotation tool f. Scalabel C. Annotation Formats: a. COCO dataset JSON files with 5 types of annotation: for object detection, keypoint detection,	

	stuff segmentation, panoptic segmentation, and image captioning. b. Pascal VOC stores annotation in XML file c. YOLO labeling format, a .txt file D. Online platform a. Amazon mechanical turk	
2. Databases for Computer Vision	A. Introduce some of the public databases with annotation and label used in classification, detection and recognition B. Classification a. MNIST b. CIFAR c. ImageNet d. PASCAL VOC e. MS COCO f. Youtube-8M g. KITTI C. Object Detection a. MS COCO b. PASCAL VOC c. KITTI D. Recognition a. Labelled Faces in the Wild b. CelebFaces E. Matching a. AT&T face	
Assignments	B. Load and visual data from KITTI dataset (autonomous driving)	

3.4. Deep Learning in Image Classification

Students should be able to :

1. Review concepts of deep learning	A. Introduce that learning feature using CNN is also part of learning in deep learning based techniques B. Review some concepts in deep learning a. CNN and describe its role in reducing the number of parameters b. Pooling c. 1x1 convolution d. Fully connected layer e. Attention layer f. Overfitting: dropout, regularization g. Loss: L1, L2, Logistic loss C. Review learning algorithm a. Backpropagation b. Review some optimizers i. SGD and Mini-batch GD ii. Adagrad iii. RMSprop iv. Adam D. Review deep learning frameworks	

	a. Tensorflow b. Pytorch	
2. Review deep learning architectures for image classification	A. Image classification a. LeNet b. Alexnet c. GoogleNet d. Network-in-Network e. ResNet f. DenseNet g. ResNeXt h. EfficientNet i. SENet (Squeeze-and-Excitation Networks) B. Techniques with low-complexity a. MobileNet b. SqueezeNet c. SqueezeNext C. Use Tensorflow or Pytorch API to demonstrate one of these algorithms (e.g. ResNet)	
Assignments	A. Implement ResNet-50 and test on ImageNet dataset or some natural images for classification.	

3.5. Deep Learning in Object Detection

Students should be able to :

1. Object Detection Algorithms	A. Introduce shortcoming of image classification algorithms in terms of localization B. Review algorithm using region proposal and existing image classification (two stage): a. RCNN C. Review some algorithms for object detection (classification + localization, together) (one stage): a. YOLO b. SSD c. SqueezeDet d. RetinaNet D. Use Tensorflow or Pytorch API to demonstrate one of these algorithms (e.g. YOLO and tinyYOLO and/or RetinaNet)	
Assignments	A. Use tiny YOLO architecture to train your face. Compare the performance with Viola's face detector.	

3.6. Deep Learning in Image Matching

Students should be able to :

1. Image matching	A. Describe image matching concept to verify whether given image is similar to the template (one example application is in biometrics for face based access control for smartphone) B. Describe Siamese CNN Network C. Use Tensorflow or Pytorch API to implement Siamese Network	

Assignment	Using AT&T face dataset design a deep-learning based face matching algorithm. Give two images of face to the network and let the network predict whether these images are the same or not.	

3.7. Module Summary

Students should be able to :

1. Summary of Object Classification and Detection	A. Understand classification, recognition, matching and detection tasks. B. Discuss on different application of Object discussion	
Project Ideas for the Module	A. Use RetinaNet to detect objects in images/video in KITTI dataset.	

4. Segmentation

4.1. Introduction to the Module

Students should be able to :

1. Image Segmentation	A. Explain segmentation problems in an image exemplified by different application areas such as medical, Self-driving cars and so on.	
2. Segmentation techniques	A. Classify Segmentation techniques in literature a. Traditional techniques which give segmentation but do not tell what object label they have. b. Deep learning based technique that relate each pixel to class label and called semantic segmentation	

4.2. Region Segmentation

Students should be able to :

1. Region segmentation techniques	A. Introduce region based segmentation B. Watershed algorithm and use OpenCV API C. Connected component D. Region based a. Region splitting algorithms b. Region merging algorithms E. Level set F. Graph cuts G. Normalized cuts H. Mean-shift I. Use OpenCV API to implement segmentation technique using Mean-shift and Graph cut	
Assignment	A. Use the Mean-shift algorithm for region segmentation for images of the KITTI dataset.	

4.3. Semantic Segmentation

Students should be able to :

1. Semantic segmentation techniques	A. Introduce semantic segmentation technique by linking each pixel to class label B. Review some of semantic segmentation techniques: a. ENet b. UNet c. Fully Convolutional Network (FCN) d. Dilated convolutions e. Mask R-CNN C. Use Mask-RCNN code to demonstrate the segmentation	Resources: https://arxiv.org/pdf/1606.02147.pdf https://www.pyimagesearch.com/2018/09/03/semantic-segmentation-with-opencv-and-deep-learning/ https://github.com/matterport/Mask_RCNN
Assignment	A. Compare performance of Dilated convolutions and Mask-RCNN to segment out images	

4.4. Module Summary

Students should be able to :

Module summary	A. Understand region segmentation and semantic segmentation	
Project for the Module	A. Use Mask-RCNN to segment out images in KITTI dataset	

5. 3D Vision

5.1. Introduction to the Module

Students should be able to :

1. Introduction	A. Understand how humans perceive 3D information B. Understand computer vision techniques for 3D vision, 3D sensors	
Applications	A. Autonomous driving B. Medical C. Cybersecurity D. Drones E. Robots F. Face matching in smartphone	

5.2. Introduction 3D Vision and Applications

Students should be able to :

1. Introduction to 3D Vision	A. Introduce 3D vision B. Introduce how human perceive 3D information C. Sensing 3D information a. 3D sensor b. Stereo correspondence	

	c. Structure from motion D. Describe following applications and why 3D vision is required: a. Autonomous driving b. Augmented reality c. Virtual reality d. Visual SLAM (Simultaneous localization and Mapping) e. Robots	

5.3. 3D Information of the Scene

Students should be able to :

1. Sensing 3D information	A. Describe 3D sensing techniques: a. Stereo camera b. Active 3D sensing i. Microsoft kinect ii. PrimeSense iii. LIDAR	
2. 3D Geometry	A. Describe important concepts in two-view geometry a. Epipolar geometry b. Stereo geometry (disparity) c. Essential matrix d. Fundamental matrix e. Eight point algorithm	
3. 3D from single image	A. Introduce concept of shape from X B. Describe techniques a. 3D from shading and texture b. 3D from focus/defocus	
4. Stereo matching (traditional)	A. Describe rectification and discuss how it eases matching B. Describe steps in stereo matching a. Matching cost computation b. Cost aggregation c. Disparity computation and optimization d. Disparity refinement C. Introduce Spare and Dense correspondence D. Describe Dense correspondence techniques: a. Local methods b. Global optimization E. Use OpenCV API to demonstrate depth map extraction	
5. Stereo matching based on deep learning	A. Give overview of public datasets: a. KITTI B. Describe some techniques for depth estimation using deep learning: a. Hourglass module b. DispNet c. CRL d. GC-Net e. EdgeStereo f. MSFNet	https://www.cs.toronto.edu/~urtasun/publications/luo_etal_cvpr16.pdf https://arxiv.org/pdf/1512.02134.pdf

Assignment	Implement stereo matching algorithm (e.g. normalized cross correlation) to estimate depth map from data in middlebury dataset	

5.4. Structure from Motion

Students should be able to :

1. Introduction	A. Introduce to recover 3D structure from image sequences	
2. Methods Structure from Motion	A. Techniques to recover structure from motion a. Triangulation b. Two frame structure from motion c. Factorization d. Bundle adjustment	

5.5. 3D Point Cloud Processing

Students should be able to :

1. Introduction	A. Describe 3D point cloud generated from LIDAR B. Introduce Point Cloud Library (PCL) C. Describe importance of LIDAR in autonomous driving and robotics D. Describe KITTI dataset in detail E. Visualize 3D point cloud in KITTI dataset	https://arxiv.org/pdf/1912.12033.pdf http://www.pointclouds.org/about/
2. Computer vision tasks on 3D point clouds	A. Describe DNN architecture for point cloud based inputs B. Describe some notable technique on 3D point clouds a. Classification: RCNet-E, PointWeb b. Object detection: 3D FCN, SECOND, PointPillars c. Object tracking: Complexor-YOLO d. Segmentation: PointNet, KP-FCNN e. Scene flow: FlowNet3D	https://arxiv.org/pdf/1912.12033.pdf https://arxiv.org/pdf/1904.07537.pdf https://arxiv.org/pdf/1903.01784.pdf https://arxiv.org/pdf/1806.01411.pdf
Assignment	A. Implement one of the object detection algorithms (e.g. PointPillars) for 3D point cloud processing.	

5.6. 3D Reconstruction

Students should be able to :

1. 3D reconstruction	A. Introduce 3D reconstruction B. Traditional 3D reconstruction from a. Surface representations b. Point based representation c. Volume representation d. Model based reconstruction	
2. Deep learning 3D	A. Describe dataset for training	https://arxiv.org/pdf/

reconstruction	a. ShapeNet B. Deep learning methods for 3D reconstruction a. 3D GAN b. OctNetFusion c. RingNet	1710.06104v2.pdf https://github.com/timzhang642/3D-Machine-Learning#3d_synthesis_dl_based
Assignment	A. Implement RingNet and experiment 3D reconstruction.	

5.7. Module Summary

Students should be able to :

Summary	A. Understand how human perceive 3D information B. Understand 3D sensor and stereo vision C. Understand algorithms to compute 3D depth map D. UnderstandMerits and demerits of stereo vision and 3D sensor for sensing depth information E. Understand real life autonomous driving data and 3D point cloud processing	
Project for the Module	A. Implement e.g. MSFNet stereo matching algorithm to estimate depth map from data in KITTI dataset.	

6. Motion & Videos

6.1. Introduction to the Module

Students should be able to :

Introduction	A. Understand need of tracking when camera or scene is motion B. Understand how motion information can lead to better signal processing e.g. Video denoising	
Applications	A. Introduction to motion in videos B. Describe the following application domains: a. Tracking object in autonomous driving b. Video stabilization c. Frame interpolation d. Video denoising	

6.2. Tracking and Motion Estimation

Students should be able to :

1. Feature tracking	A. Describe sparse feature tracking method B. Describe Kanade-Lucas-Tomasi (KLT tracking) a. Describe corner feature detection b. Describe tracking of features and brightness	

	constraint C. Describe KLT tracking in pyramid D. Use OpenCV API to demonstrate tracking of feature in video	
2. Dense motion estimation	A. Describe motion models B. Describe dense motion estimation and contrast with spare feature tracking C. Describe optical flow technique for dense motion estimation D. Use OpenCV API to implement dense motion estimation	
3. Object or window tracking (traditional)	A. Introduce the need of object tracking e.g. pedestrian, car, etc. B. Describe HOG and color histogram for pedestrian tracking C. Describe Mean Shift algorithm D. Describe Camshift algorithm E. Use OpenCV API to demonstrate object tracking in a video	
4. Deep learning based object tracking	A. Introduce the CNN and Siamese deep network for object tracking B. Describe multi-object tracking in the scene C. Describe the techniques in the literature a. MDNet and GOTURN b. SiamFC c. SiamRPN d. SiamRPN++ D. Use Tensorflow or PyTorch library to implement object tracking using one of these techniques	https://www.robots.ox.ac.uk/~luca/siamese-fc.html
5. Filtering based tracking	A. Introduce tracking as inference B. The Kalman filter C. Particle filter D. Use OpenCV API to demonstrate to use the Kalman filter to track the window/object (e.g.car) in the video	
Assignment	A. Implement Camshift algorithm to track object some dataset e.g. KITTI.	

6.3. Dense motion Estimation

Students should be able to :

1. Optical flow (traditional)	A. Describe dense motion estimate B. Describe steps in optical flow	
2. Optical flow based on deep learning	A. Describes technique based on DL: a. FlowNet b. SceneFlowNet	
Assignment	A. Implement SceneFlowNet and understand the architectural details.	

6.4. Module Summary

Students should be able to :

Module Summary	A. Understand object tracking and how it is different from object detection B. Understand traditional and deep learning based tracking algorithms	
Project for the Module	A. Implement SiamFC and use it to track objects in the KITTI dataset.	

7. Generating Synthetic Images

7.1. Introduction to the Modul

Students should be able to :

1. Introduction to Synthetic Image Generation	A. Applications of synthetic image generation from the distribution a. Generate more training data b. Image/video restoration B. Distinguish with computer graphics for: shading, texture mapping, shadows, reflection, etc.	

7.2. Image-based rendering

Students should be able to :

1. Image-based rendering	A. Introduce image based rendering B. Describe image based rendering: a. View interpolation b. Layered depth images c. Light fields d. Environment mattes e. Video-based rendering	

7.3. Generative Adversarial Network

Students should be able to :

1. GAN	A. Recall and describe generative and discriminative models and explain how GAN works a. Describe role of generator and discriminator B. Describe conditional GAN C. Describe DCGAN a. Describe vector operation on visual objects D. Explain the difficulties in training GAN	https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf https://arxiv.org/pdf/1511.06434.pdf

	a. Vanishing gradient b. Mode collapse c. Failure to converge E. Remedies for training GAN a. Wasserstein loss b. Unrolled GANs c. Adding noise to discriminator inputs d. Penalizing discriminator weights	https://developers.google.com/machine-learning/gan/problems https://arxiv.org/pdf/1701.04862.pdf https://arxiv.org/pdf/1705.09367.pdf
2. Applications of GAN	A. Discuss on list applications of GAN a. Frame interpolation b. Super-resolution c. Blur removal d. Image to image translation e. Text to image generation f. Inpainting g. Generating training data B. Describe following applications in detail: a. Image to Image translation	https://arxiv.org/pdf/1411.1784.pdf https://phillipi.github.io/pix2pix/

7.4. Module Summary

Students should be able to :

Module Summary	A. Understand image rendering techniques and their applications B. Understand GAN and difficulties surrounding GAN while training and remedies C. Understand GAN to generate images	
Project for the Module	A. Implement cGAN in Keras and generate cartoons with hair and eyes with different colors.	

8. Deep Learning in Image and Video Processing

8.1. Introduction to the Module

Students should be able to :

1. Introduction to the Module	A. Understand the role of deep learning on image and video processing. B. Understand different deep learning algorithms for super-resolution, deblurring and frame interpolation.	

8.2. Superresolution

Students should be able to :

1. Introduction	A. Introduce super-resolution B. Discuss GAN for Super-resolution C. Perceptual loss and VGG features D. Effect of L1 and L2 loss in training	
2. Super-resolution	A. Discuss public dataset for super-resolution: a. DIV2K b. Set5, Set14 c. B100 d. Urban100 e. Vimeo90k f. Vid4 B. Deep learning algorithms for super-resolution a. SRCNN b. ESPCN and Sub-pixel CNN c. DRCN d. EDSR e. EnhanceNet f. RBPN and DBPN C. GAN based algorithms: a. SRGAN b. FCGAN c. TecoGAN	
Assignment	A. Implement EDSR and use computer super-resolution of some real-life images. Compare the visual quality with bicubic interpolation.	

8.3. De-blurring

Students should be able to :

1. Introduction	A. Introduce de-blurring and role of deep learning B. Perceptual loss and VGG features C. Adversarial loss D. Effect of L1 and L2 loss in training	
2. De-blurring	B. Discuss public dataset for de-blurring: a. GOPRO b. Kohler dataset C. Deep learning for deblurring a. DeepDeblur b. SRN D. GAN a. DeblurGAN	
Assignment	A. Visualize images in GOPRO dataset for deblurring and use DeepDeblur to recover sharp image.	

8.4. Frame interpolation

Students should be able to :

1. Introduction	A. Introduce frame-interpolation	

	B. Discuss importance of frame interpolation C. Discuss role of CNN with features in frame interpolation	
2. Frame interpolation	A. Datasets a. Vid4 B. Deep learning based algorithms a. SepConv b. SuperSlomo c. DAIN C. GAN based algorithms a. FINNiGAN b. FIGAN	
Assignment	Use SuperSlomo to create intermediate video frames.	

8.5. Module Summary

Students should be able to :

Summary	A. Understand the role of CNN and GAN on image and video restoration tasks B. Understand merits and demerits of L1 and L2 loss and perceptual and adversarial loss.	
Project	A. Use DeepDeBlur orDeblurGan to de-blur the image in the KITTI dataset and apply RetinaNet to detect objects. Observe whether there is improvement in object detection.	

9. Computational Photography and Imaging

9.1. Introduction to the Module

Students should be able to :

Introduction	A. Differentiate between computational photography and computer vision in general B. Describe flash/no-flash images C. Describe multi-exposure images with different integration times D. Describe applications: a. High dynamic range b. Super-resolution 4K TV c. Lytro adjusting depth of field (refocus) d. L16 camera	

9.2. High dynamic range (HDR)

Students should be able to :

	A. Give introduction to high dynamic range (HDR)	

1. Introduction	B. Give some example pictures of HDR images	
2. Traditional technique to create HDR	A. Discuss tone mapping B. Use OpenCV API to do tone mapping C. Introduce HDR file format D. Discuss HDR reconstruction from multiple exposures	http://static.googleusercontent.com/media/hdrplusdata.org/en//hdrplus.pdf
3. Deep learning for HDR	A. Discuss dataset for HDR B. Deep learning technique for HDR from single exposure C. Deep learning technique for HDR from multiple exposures	https://github.com/gabrieleilertsen/hdrcnn

9.3. Computational Processing

Students should be able to :

1. Introduction	A. Describe image recovery from multiple images e.g. different exposures, illumination, etc.	
2. Blur removal	A. Describe optical blur and motion blur (or camera shake) B. Visualize images at different exposure values (images at short exposure being noisy and long exposure with camera shake) C. Describe deconvolution of camera shake image with long exposure with image with short exposure as prior D. Burst image deblurring	https://dl.acm.org/doi/10.1145/1275808.1276379 http://openaccess.thecvf.com/content_ECCV_2018/papers/Miika_Aittala_Burst_Image_Deblurring_ECCV_2018_paper.pdf

9.4. Computational Imaging/Optics

Students should be able to :

1. Concepts on computational imaging	A. Describe computational imaging B. Describe difficulties of inverse filtering (de-convolution) C. Describe coded aperture (and flutter shutter) for motion blur removal D. Describe plenoptic camera and digital refocus	

9.5. Computational Sensor

Students should be able to :

1. Concepts on computation sensor	A. Event camera or DVS sensor B. Fluttered shutter a. Explain the application of motion deblurring	https://www.zora.uzh.ch/id/eprint/17629/1/Lichtsteiner_Latency_V.pdf https://inivation.com/dvs/ http://web.media.mit.edu/~raskar/deblur/

		http://web.media.mit.edu/~raskar/

9.6. Module Summary

Students should be able to :

Module Summary	A. Discuss the role of hardware on simplifying the computation for deblurring and tracking B. Combine multiple frames to create more HDR like image C. Understand how some sensors can be used for tracking such as DVS and how they are different from frame cameras.	
Project for the Module	A. Implement a burst imaging technique for deblurring and to create HDR image.	

10. Trustworthy Computer Vision

10.1. Introduction to the Module

Students should be able to :

Introduction	A. Introduce what is explainable AI and trustworthy AI B. Describe Clever Hans phenomenon C. Describe some adversarial samples and method to generate adversarial samples through optimization D. Explain deep models for image classification, matching, etc. are black box and difficult to interpret E. Explain some problems: change of one pixel classify pandas as giboon	http://iphome.hhi.de/samek/ https://openai.com/blog/adversarial-example-research/ https://arxiv.org/abs/1712.07107 https://medium.com/@ml.at.berkeley/tricking-neural-networks-create-your-own-adversarial-examples-a61eb7620fd8

10.2. Explainable Computer Vision

1. Explanation methods	A. Introduce trustworthy AI and explainable AI B. Describe Clever Hans phenomenon C. Explain deep models for image classification, matching, etc. are black box and difficult to interpret the output D. Model agnostic a. LIME (https://github.com/marcotcr/lime)	https://christophm.github.io/interpretable-ml-book/intro.html https://github.com/kundajelab/deeplift

	b. SHAP (https://github.com/slundberg/shap) c. ANCHOR E. Deep model specific a. Grad-CAM (https://arxiv.org/pdf/1610.02391.pdf) b. Deep LIFT c. LRP F. Understand merits and demerits of the current explanation techniques (Interpretable Machine Learning)	http://iphome.hhi.de/samek/
Assignment	A. Use SHAP explanation to explain output of models for: a. Sentiment analysis. Identify keywords used for prediction and discuss whether these words look relevant.	

10.3. Robustness and adversarial attacks

1. Robustness and adversarial attacks	A. Introduce the adversarial attacks B. Describe some adversarial samples and method to generate adversarial samples through optimization C. Explain some problems: adding noise classify a panda as a giboon D. Describe some papers: a. Explaining and Harnessing Adversarial Examples b. One Pixel Attack for Fooling Deep Neural Networks c. Detecting Adversarial Samples from Artifacts d. Face Anti-Spoofing Using Patch and Depth-Based CNNs e. A Dataset and Benchmark for Large-scale Multi-modal Face Anti-spoofing E. Verifying Properties of Binarized Deep Neural Networks	https://openai.com/blog/adversarial-example-research/ https://medium.com/@ml.at.berkeley/tricking-neural-networks-create-your-own-adversarial-examples-a61eb7620fd8 https://arxiv.org/abs/1712.07107 https://arxiv.org/pdf/1412.6572.pdf https://arxiv.org/pdf/1710.08864.pdf https://arxiv.org/pdf/1703.00410.pdf
Assignment	A. Generate adversarial samples by modifying one pixel and reproduce results of one pixel attack.	

10.4. Module Summary

Students should be able to :

Module Summary	A. Able to debug ML models with explainability and verify whether operating in the expected way. B. Understand tool and how to generate explanation for given instance of input	

	C. Able to verify models and whether operating correctly D. Incorporate adversarial concepts to strengthen models against it.	
Project for the Module	A. Show the heatmap of the region in the image used for classification. Use this tool to verify your object classification algorithm.	

Glossary of command words

This glossary table below is a guide to develop learning outcomes for each chapter that would help to develop, teach and learn content as well as for evaluation.

Command words	**AO**	**Indication (What it means)**	**Use Cases**
Define,	1	Formal statement/phrases of definition	equations, terms
What is meant by	1	Definition + Significance or context of the term	
Describe	2	State the main points, with diagrams, examples or so on	Phenomena, experiment, observation
Explain	2	Reasoning with reference of theory applied	Relationship, cause-effects,
State	1	Express in concise words, without supporting arguments	Theorem, law, fact, value without calculation
List	1	List down number of points without elaboration	
Exemplify	1,2	Provide examples	
Discuss	5	Critical account on the topic	critique,
Deduce, solve, Predict	4	Produce the answer through logical connections than just by recall	
Suggest, propose	4	applying knowledge to a new situation or when there is no unique idea	
Calculate, find out, workout, carry out	3	get a numerical from given data, some value, through some work	Workout to calculate a value
Determine	3	Quantity calculated with certainty	Magnitude, scale ,
Show	3	Derive result through a structured explicit evidence	
Justify, support	4	Support a case with evidence/arguments	
Verify, prove	4	Confirm that a given statement/result is true	
Estimate	4,5	Reasoned order of magnitude for the quantity	
Sketch	3,4	Make freehand sketch drawing/curve with key features	Diagrams, graphs, figures
Compare	4	Provide similarities and differences	

Recognise, identify, name, select	1	identify from having encountered them before	
Implement, code	3,6		
Create, design, construct,	6		
understand			
appreciate			