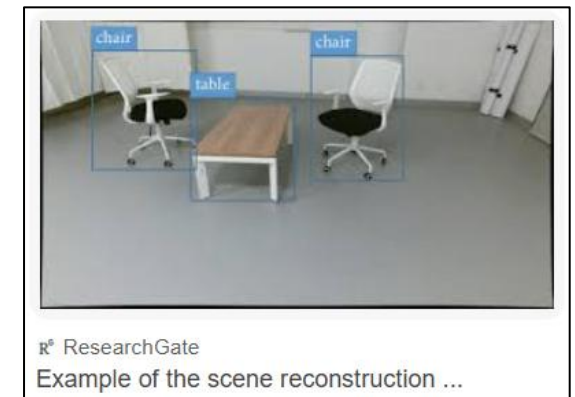
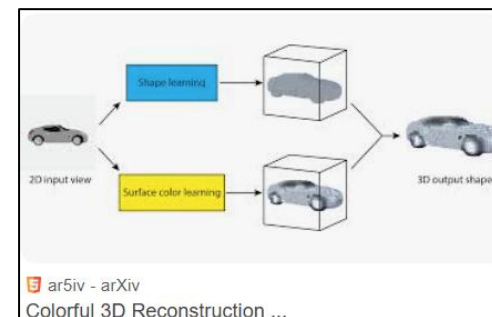
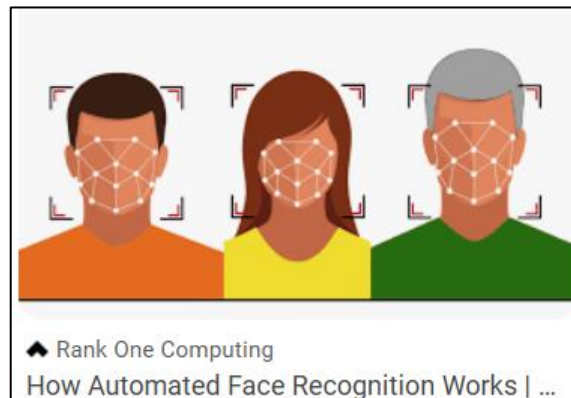
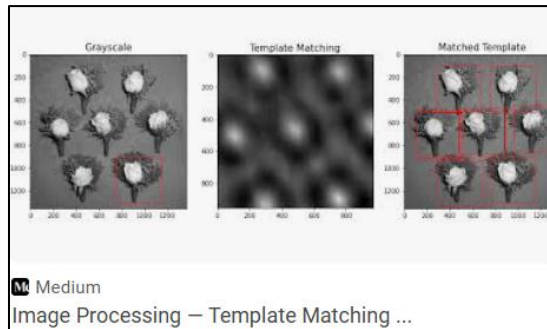


Module 1

Computer Vision

CV is an area of AI that makes computers identify and analyze images and videos.





This module gives us a glimpse of:

- Brief history of CV
- Cameras, Sensors and Digital Images
- Computer Power
- Machine Learning
- Big Data CV Applications



Assignment: A01 – History of Computer Vision

Module 2



CV Cameras and Censors:

- Camera Specification – Field of View, Frame Rate, Resolution
- Stereo and Thermal; Depth and RGB
- Radar and LiDAR Sensors
- Setting the camera up and calibrating it
- Discussing the benefit and the drawback



Module 3

CV Libraries, Frameworks, and IDEs:

Here, we learned about how python (its libraries and frameworks) works well with Jupyter Notebook and Labs. There are some other resources like ML Libraries (OpenCV, TensorFlow, Pytorch) and ML Frameworks like Keras, Fastai and Caffe. Aside from these, there are also LDE (Visual code Studio); HPC (TACC); CR&CP (Github, Hugging Face); CBDE (Google Colab, Sagemaker Studio Lab, Azure AI). Then come the AI Code Assistants like Gemini, Github Copilot, Microsoft Copilot, Amazon CodeWhisperer and Devin Ai.

Assignments – A03 Understanding the tools of the trade (Jupyter Notebook & Labs)
L03 Practical Exploration of Github and Jupyter

Module 4

Introduction to Image Processing

- Image processing is the process of handling and evaluating images to get relevant informations to improve the quality.
- Techniques for enhancing images
- Formats and representations of digital images
- Pillow and OpenCV

Assignment – A04 Image Processing Adventure Quest

Module 5

Machine Learning for CV

Here, we have the chance to learn the ML vocabulary like feature, label, parameter, dimensionality, feature training and model training. The difference of supervised, unsupervised and reinforcement learning is clarified. Major components to ML like data, compute and algorithms is also explained. MLDQs is also a part of machine learning.

Assignment – L05 Image Classification (Reflective Journal)

Introduction to Neural Networks

- NN has been inspired by the construction and the purpose of the human brain.
- It contains nodes called neurons connected to one another and arranged in layers.
- Deep Learning, Traditional Learning and Machine Learning
- The Artificial Neuron
- Training Neural Networks
- The Learning Process – starts with the initialization, then the feeding forward, followed by calculating the loss, next is backpropagation, weight update, repetition and last is evaluation.
- Some of these vocabulary words are also parts of the process – gradient descent, output function, model evaluation, regularization and optimization.
- Hyperparameter tuning
- Neural network architecture

Assignment: Lab 6 TensorFlow Playground (Chihuahua or Muffin? – Reflective Journal)

Module 7

Convolutional Neural Networks - CNN

- The focal elements here are the Input Layer, Convolutional Layers (for feature extraction), Fully Connected Layers (for classification) and Output Prediction.
- The center elements of CNN are Convolutional layers, Pooling layers, Fully Connected layers.
- CV Datasets – MNIST, CIFAR-10, Open Images, Places, ImageNet, Caltech 101, Caltech 256, Vehicle make and model rec. dataset

Module 8

Transfer Learning & CNN Core Architectures

- Pre-trained Models – VGG, ResNet, Inception, EfficientNet and Vision Transformers
- The special toolkits are – NVIDIA, Pytorch, TenseFlow and MXNET GluonCV.
- Evolution of CNNs-
 - LeNet-5 (1998)
 - AlexNet (2012)
 - VGGNet (2014)
 - GoodLeNet/Inception (2014)
 - ResNet (2015)

Assignments; A07 Manual CNN; L07 Basic CNN

Module 9

Building Bridge: From CNN Basics to Object Detection Architectures

Here, we gain an understanding of advanced CNN architectures, General CNN architectures and the Region-Based Detection (R-CNN, SPP-Net, Fast R-CNN, Faster R-CNN). We also touched the topics about Bounding boxes, object detection and image classification.

Assignment: L09 Object Detection using TensorFlow and COCO Dataset

Midterm Project: Object Detection Challenge

Module 10

Mastering CV in Video Fun-damentals

- Video
- The important parts – resolution, compression, frame rate, color space
- New improvement in video analysis – Diffusion models, transformers, state-space models

Module 11

Generative Models in Computer Vision

In this field, we have the generative AI like Gans, Transformers, VAEs, SSMs and Diffusion models. When it comes to generalist and foundation models, we have resources like DALL-E, BERT, Stable Diffusion and GPT.

Module 12

Autonomous Systems in CV

The AI agents are our helpers. They guide us when we ask for help in how to be successful in what we do. They are reactive, deliberative and hybrid. They can be found in robotic agents and software.

Assignment: Capstone Project (Train an AI agent to play Flappy Bird)

References:	
Ar5iv-arxiv	Jupyter Notebook
Colab	Medium
Gemini API	Rank One Computing
Github	ResearchGate
ITAI 1378 2024 Modules 1 to 12.pptx	

<https://github.com/idv713/jupyter-exploration>