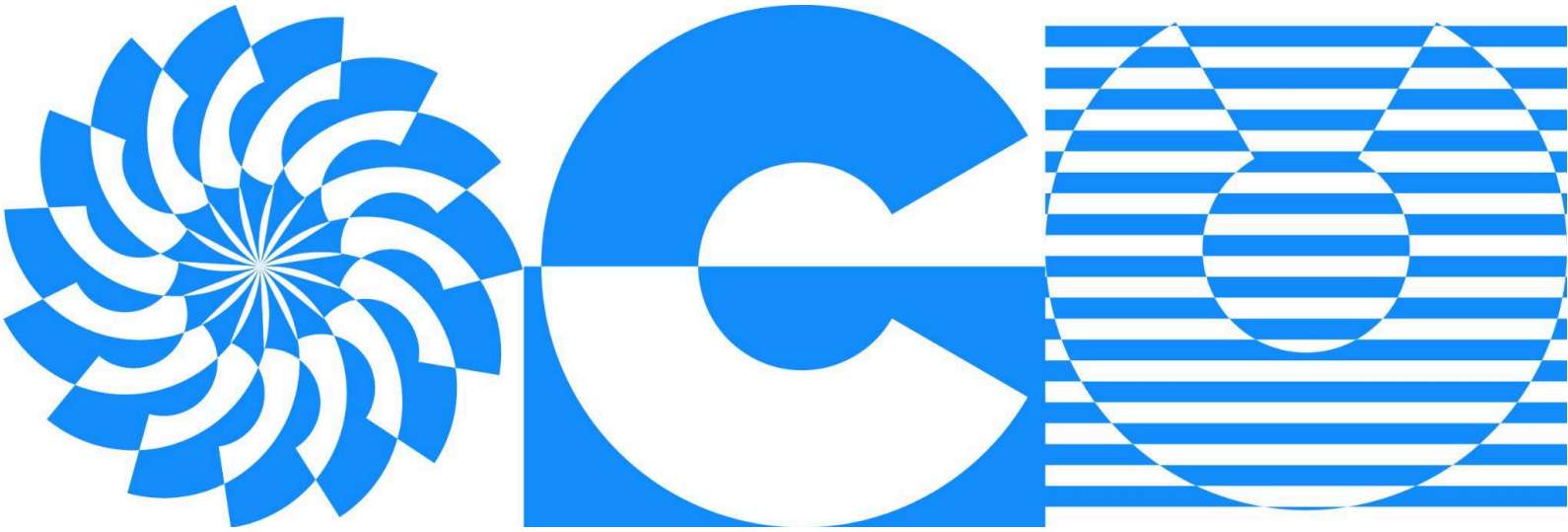
A place where legacy creates the future.



TXAP

Advanced Vision Applications with Deep Learning & Transformers

Detailed Curriculum

OpenCV University

Index

Introduction

Welcome

- Welcome
- Introduce Yourself!

Course Logistics

- How is the Course Organized?
- Grading Policy and Certification
- How to Use Online Labs
- Where can I Download the Code?
- Quiz Instructions

Setting Up Deep Learning Tools

- Introduction to Colab
- Introduction to Kaggle Kernel

Module 1: Neural Networks and Classification

Download Code

- Download Code

Introduction to Neural Networks

- What is a Neural Network?
- Deep Learning Frameworks
- How Does a Neural Network Learn?

Introduction to PyTorch

- What is PyTorch and Why Choose PyTorch?
- PyTorch Basics and Tensor Operators
- Training Neural Networks: Overview

- Binary Classification with PyTorch
- Dataloaders in PyTorch: Classification

Feedforward Neural Networks

- Why do We Need Hidden Layers?
- Training a Neural Network using Backpropagation
- Image Classification using MLP
- Example: Image Classification using MLP

Convolutional Neural Networks

- The Convolution Operation
- Example: Image Classification using CNNs

Logging using MLOps Tools

- Introduction to Weights & Biases

Transfer Learning and Fine-Tuning using Pre-Trained Models

- Transfer Learning Notebook
- Train Image Classifier to Fine-Tune on a Difficult Dataset

Vision Transformers

- Introduction to ViT
- ViT Attention Map
- Fine-Tuning ViT for Birds Classification

Assignment 1

- Assignment 1: Binary Classification (Graded)

Quiz 1

- Quiz 1 (Graded)

Module 2: Object Detection

Download Code

- Download Code

Introduction

- Object Detection Overview
- Traditional Object Detection Pipeline

Single Stage Detectors

- SSD
- RetinaNet

Object Detection using YOLO

- Introduction to YOLO
- Object Detection Inference
- Fine-Tuning for Aerial Images
- Tiled Object Detection using YOLOv8

Detection Transformers

- Object Detection using RTDETR

Case Study - Global Wheat Challenge

- Global Wheat Detection Part 1 - Training
- Global Wheat Detection Part 2 - Practical Improvements

Assignment 2

- Assignment 2: Annotation Conversion (Graded)

Quiz 2

- Quiz 2 (Graded)

Project 1

- Object Detection (Graded)

Module 3: Text Detection & Recognition (OCR)

Download Code

- Download Code

Introduction

- Introduction

Graphic Text Recognition using Tesseract

- What is Tesseract?
- Notebook: Introduction to OCR using Tesseract
- Notebook: Tesseract OCR Failure Cases
- Notebook: Improving Tesseract OCR Failures

Transformer OCR (TrOCR)

- Introduction to TrOCR
- TrOCR Inference with Cropped Image
- TrOCR Inference with Text Detection
- Fine-Tuning using TrOCR on Captcha

Application - ALPR

- Introduction to ALPR
- YOLOv10 License Plate Detection Training
- Fine-Tuning TrOCR on License Plate Text
- ALPR Inference - Combined Detection + OCR

Assignment 3

- Assignment 3: TrOCR Invoice (Graded)

Quiz 3

- Quiz 3 (Graded)

Project 2-

- TrOCR (Graded)

Module 4: Segmentation

Download Code

- Download Code

Introduction

- Introduction to Segmentation

Deep Learning-Based Segmentation Models

- Introduction to Segmentation using Torchvision Models

Transformer-Based Segmentation Models

- Aerial View Segmentation using SegFormer

SAM (Segment Anything Model)

- Introduction to SAM
- Automatic Person Segmentation with YOLOv11 + SAM2
- Advanced Segmentation Techniques with SAM2

Segmentation using Custom Backbone

- DINO UNet Road Segmentation

Quiz 4

- Quiz 4 (Graded)

Project 3

- Segmentation (Graded)

Module 5: Tracking

Download Code

- Download Code

Introduction

- Overview
- Object Tracking Algorithms

Ultralytics Tracking

- Tracking using Ultralytics YOLOv8
- SeaDrone Dataset Detection Fine-Tuning and Tracking
- YOLO11 Grocery Cart Tracking and Counting

Multi-Camera Tracking

- Multi-Camera Tracking using OpenVINO

Re-Identification

- Person ReID Fine-Tuning

Transformer-Based Tracker

- Introduction to Point Tracking using CoTracker3

Quiz

- Quiz 5 (Graded)

Module 6: Keypoint Estimation

Download Code

- Download Code

Introduction

- Introduction to Pose Estimation and Landmark Detection

Exercise Analysis using Keypoints

- Squats Analysis Pipeline Overview
- Squats Analysis Code Explanation

Keypoint Fine Tuning using YOLOv8

- Facial Keypoint Fine Tuning using YOLOv8

Applications of Keypoint

- Warping a Triangle
- Delaunay Triangulation
- Face Alignment
- Face Averaging
- Face Morphing
- Bug Eyes
- Face Swap
- Beard Filter
- Aging Filter

Assignment & Quiz

- Assignment 4: Smile Detection (Graded)
- Quiz 6 (Graded)

Module 7: Face Recognition and Applications

Download Code

- Download Code

Creating an Automated Attendance System using AWS

- Introduction to Automated Attendance System
- Introduction to AWS Rekognition
- How to Setup AWS with AWS Rekognition
- Registration using AWS Rekognition
- Deregistration using AWS Rekognition
- Attendance Using AWS Rekognition
- Gradio App Setup Instruction

Assignment & Quiz

- Assignment 5: Doppelganger (Graded)
- Quiz 7 (Graded)

Bonus Module

Download Code

- Download Code

CllP

- Zero Short Image Classification with CLIP
- Image Captioning and VQA with Moondream2

Summary of Assessments

- **Total Quizzes:** 7
- **Total Assignments:** 5
- **Total Projects:** 4

Explore Other Courses