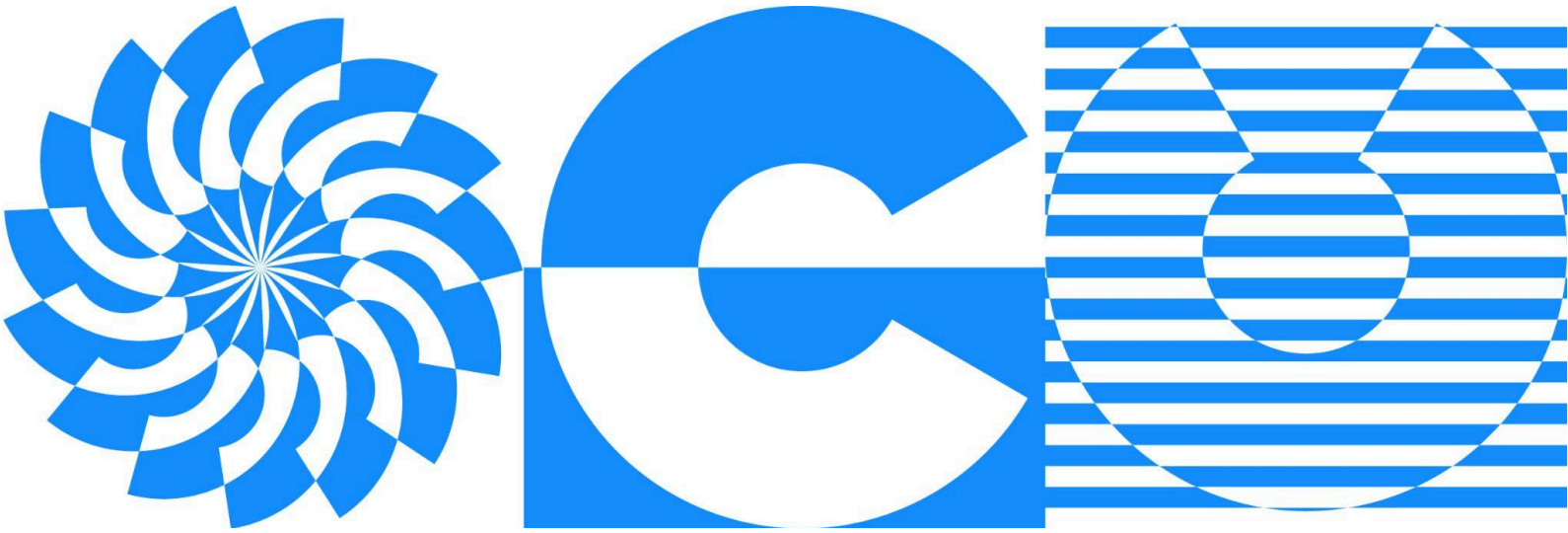


A place where legacy creates the future.



TxAP

# **Advanced Vision Applications with**

# **Deep Learning & Transformers**

Detailed Curriculum



# Index

1. **Introduction**
  2. **Module 1: Neural Networks and Classification**
  3. **Module 2: Object Detection**
  4. **Module 3: Text Detection & Recognition (OCR)**
  5. **Module 4: Segmentation**
  6. **Module 5: Tracking**
  7. **Module 6: Keypoint Estimation**
  8. **Module 7: Face Recognition and Applications**
  9. **Module 8: Vision-Language Models (VLMs)**
-

# Introduction

## Welcome

## Course Logistics

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

## Setting Up Deep Learning Tools

- Introduction to Colab
  - Introduction to Kaggle Kernel
- 

# Module 1: Neural Networks and Classification

## 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 with CNN

## **Quiz 1**

- Quiz 1
-

# Module 2: Object Detection

## Introduction

- Object Detection Overview
- Traditional Object Detection Pipeline

## Object Detection using YOLO

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

## Single-Stage Detectors

- SSD
- RetinaNet

## 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

## Quiz 2

- Quiz 2

## Project 1

- Object Detection
- 

## Module 3: Text Detection & Recognition (OCR)

### 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

### Quiz 3

- Quiz 3

## Project 2

- TrOCR
- 

# Module 4: Segmentation

## 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

## Project 3

- Segmentation
-

# Module 5: Tracking

## 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
- 

# Module 6: Keypoint Estimation

## 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
- Quiz 6

---

# Module 7: Face Recognition and Applications

## 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

## **Custom Face Verification System**

- Introduction and Setup
- Face Verification Pipeline

## **Assignment & Quiz**

- Assignment 5: Doppelganger
  - Quiz 7
- 

# **Module 8: Vision Language Models(VLM)**

## **Zero Shot Image Classification with CLIP**

- Code Explanation

## **Moondream2 and Paligemma2-Mix**

- Image Captioning and VQA with Moondream2
- Introduction to Paligemma2 Mix

## **Math Equations to Latex: Fine-Tuning Gemma 3**

- Fine-Tuning Gemma 3 Notebook
- Math Equation to Latex Gemma 3 Inferencing

## **Video Comprehension**

- Video Summarization Using Qwen-2.5VL
- Content Moderation using CLIP and Gemini

## Video RAG

- [IntroductiontoVideoRAG](#)
  - [VideoRAGPipelineCodeExplanation](#)
- 

## Module 9 – Deployment

### Introduction

- [Deploying Language Models on Local and Cloud](#)

### Get Familiar with Local Quickstart Tools

- [The Existential Problems in Autoregressive Inference](#)

### Why Model Serving Engines Exist?

- [The Existential Problems in Autoregressive Inference](#)

### Introduction to vLLM: The Throughput King

- [Model Serving Engines Deep Dive](#)

### Quiz

- [Quiz 9](#)

## Summary of Assessments

- **Total Projects: 3**
- **Total Assignments: 5**
- **Total Quizzes: 9**

[Explore Other Courses](#)