# Homework 4 Spring 2025

## Applied Computer Vision (CMU-Africa)

**Release Date: Thursday, March 13$^{th}$, 2025 CAT**
DUE: **Friday March 31$^{st}$, 2025, 11:59 PM CAT**

# Object Detection and Image Segmentation

In this assignment, you will use transfer learning to fine-tune an object detection model. Additionally, you will experiment with various image segmentation models to analyze their effectiveness in different scenarios.
For the image segmentation task, we will apply segmentation techniques to count objects similar to what we did in Homework 2, where we relied on traditional image processing techniques. This will allow you to compare classical methods with deep learning-based segmentation.

1. **Collaboration policy:**Students are encouraged to work in groups but each student must submit their own work. If you work as a group, include the names of your collaborators in your write up. The code should NOT be shared or copied. Please DO NOT use external code unless permitted. Plagiarism is strongly prohibited and may lead to failure of this course.

2. **Start early:** This is a much bigger assignment than assignment 1, 2 & 3.

3. **Getting Help**: If you have any questions, please look at Piazza first. Other students may have encountered the same problem, and it may be solved already. If not, post your question on the discussion board. Teaching staff will respond as soon as possible.

4. **Reporting:** Your write-up should mainly consist of three parts, your answers to theory questions, resulting images of each step, and the discussions for experiments. Please note that we DO NOT accept handwritten scans for your write-up in this assignment. Please type your answers to theory questions and discussions for experiments electronically.

5. File paths: Please make sure that any file paths that you use are relative and not absolute. Not cv2.imread('/name/Documents/subdirectory/hw4/data/xyz.jpg') but cv2.imread('../data/xyz.jpg').

6. **Submitting your work:**

   - You will need to fill this form
   - Dataset should be public on Github or Google Drive
   - A notebook containing codes for Part 1 and Part 2
   - A report on your procedure, outputs and reflections.

## Objectives

By the end of this assignment, you will:

1. Train an object detection model using transfer learning on a custom dataset.
2. Experiment with different segmentation models and evaluate their effectiveness.
3. Use segmentation for object counting, comparing results with traditional image processing techniques.
4. Analyze and compare performance using appropriate metrics such as Intersection over Union (IoU) and mean Average Precision (mAP).
5. Understand the trade-offs between object detection and segmentation for various applications.

# Part 1: Object Detection

## Step 1: Group Formation

- Form groups of 5 students.
- Choose one object that the entire group will detect (e.g., book, cup, chair).
- Each group member must use a different object detection model, such as:
  - YOLO
  - Faster R-CNN
  - SSD (Single Shot Detector)
  - EfficientDet
  - RetinaNet

  We do not expect you to build this models from scratch but use frameworks like Pytorch and TensorFlow. 5 bonus points to anyone who uses add objects from another group to their model.

## Step 2: Dataset Collection & Annotation (COCO Format) (Due Sunday 16th March 2025)

- Each Student must capture 20-30 images of the chosen object using their smartphone.
- The dataset should contain at least 100-150 images per group
- Label images using COCO format annotations using tools like Labelme, CVAT, or RoboFlow.

### Step 3:Modify the Object Detection Model

Each group member must modify their chosen model to detect the selected object. Follow these steps:

1. Download & Load Pretrained Weights.
2. Prepare Custom Dataset (COCO Format Required)
3. Train the Model on Custom Data
4. Fine-Tune Hyperparameters.
    - Adjust learning rate, batch size, and epochs based on dataset size.
    - Use data augmentation to increase model robustness.
5. Evaluate Model Performance
    - Compute mAP (mean Average Precision) and IoU (Intersection over Union).

### Step 4: Model Comparison & Performance Analysis

Each group member should compare their model's performance using:

- mAP scores
- Precision & Recall
- Inference speed (FPS)
- Detection quality (Visual Analysis)

Final Group Discussion:

- Which model performed best?
- Trade-offs between accuracy and speed.
- Which model generalizes better?

## Part 2: Image Segmentation

### Select a Segmentation Model

Use at least 3 different segmentation models which could be:

- Mask R-CNN
- DeepLabV3
- U-Net
- SegFormer
- Segment Anything Model (SAM)
- etc.

## Perform Image Segmentation

1. Select an image containing multiple instances of the object they worked on in Homework 2 (Object counting).
2. Segment the image to isolate the objects using the chosen model.
3. Visualize the segmented objects with different colors or masks.

## Count Objects

Apply connected component analysis or contour detection to count the segmented objects. Report on the counts of each model on the images provided in previous assignment.