Deeploy CV Project

Assignment : 2

Name: Vaibhav Itauriya

Roll Number: 231115

ID : 223

Branch: ME

Date of submission: 20/12/2024

GitHub Repository of this Project:

Answer 1

This question helped me explore two popular object detection models, YOLO and SSD, and understand their architectures and capabilities. Implementing these models taught me to classify and detect multiple objects in real-time video. It was fascinating to see how different models perform in similar scenarios and identify their strengths and weaknesses. This exercise sparked my curiosity about object detection and its real-world applications. Using the iou threshold and score threshold, I have tried this to identify different things.

YOLO

For YOLO, I used the `YOLOv5s` model, which is optimised for speed and accuracy.

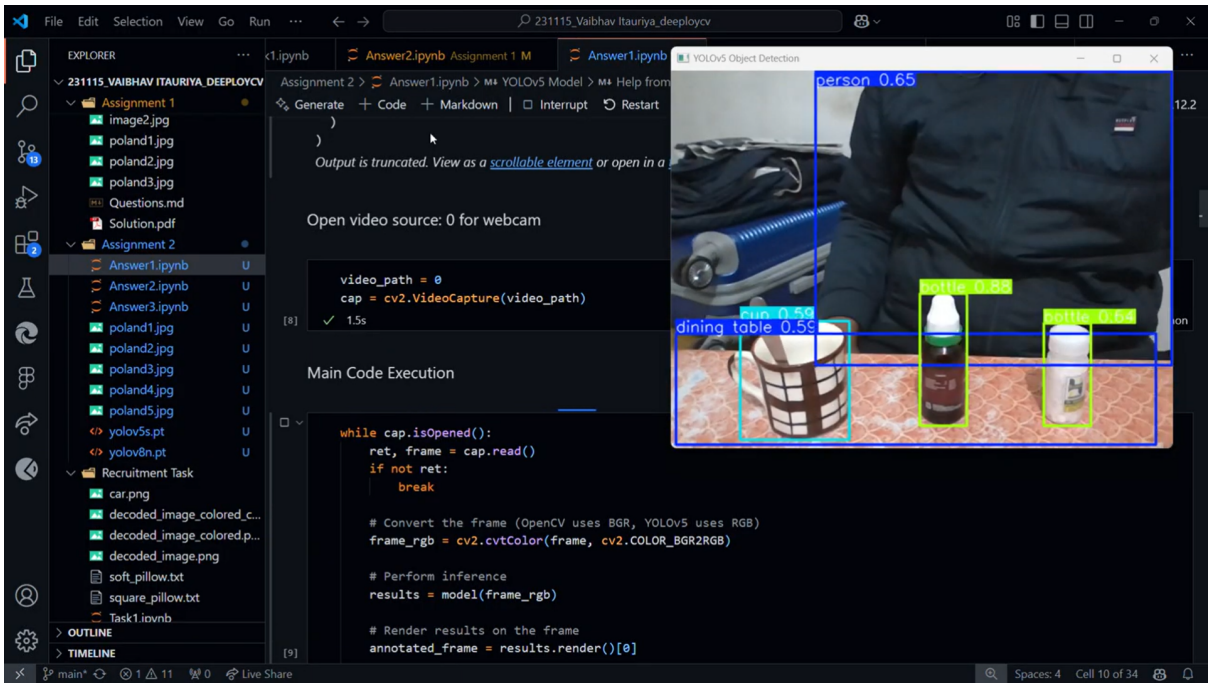

Figure 1: YOLOv5s Model

YOLOv5s Model Working Video

- Detected most objects in the frame quickly.
- Performed well on small and overlapping objects due to its grid-based approach.
- Higher FPS, making it suitable for real-time applications.

SSD

For SSD, I implemented the `SSD300 with VGG16 backbone`, which balances performance with computational efficiency.

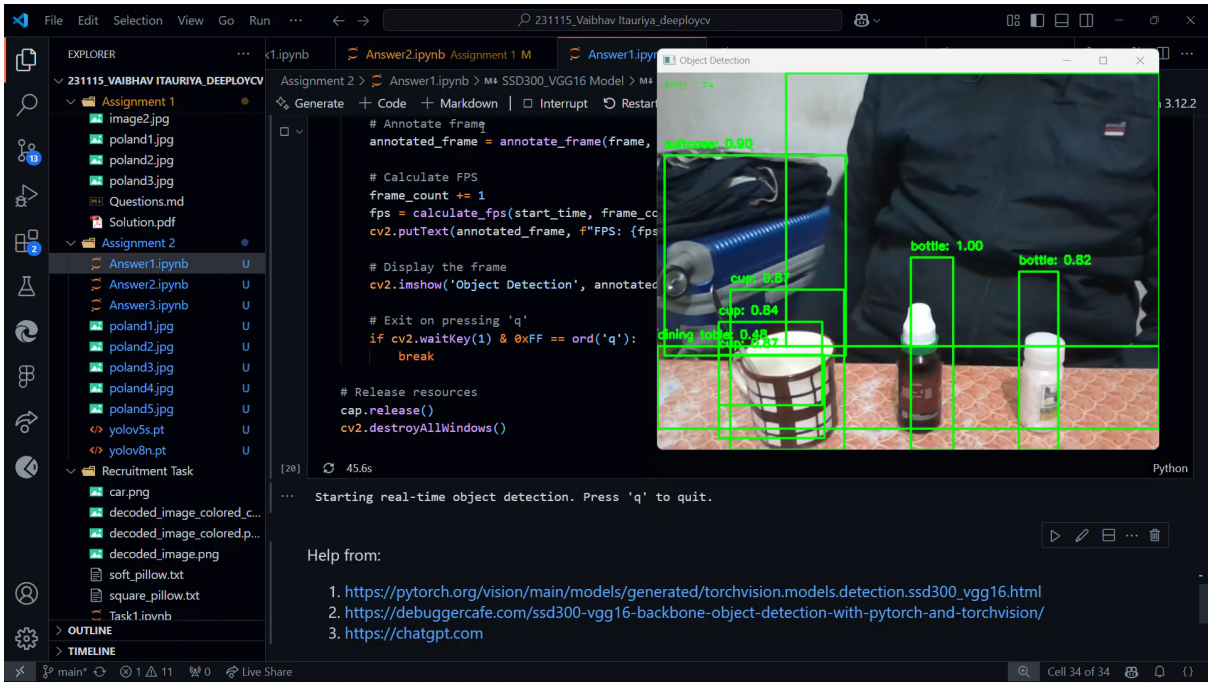

Figure 2: SSD Model
SSD300_VG16 Model Working Video

- Accurate on larger objects but struggled with smaller or overlapping objects.
- Slightly slower than YOLO due to its architecture and bounding box predictions.
- Detected items, but it was too choppy, and the delayed response was there.

Differences Noticed

In side-by-side comparison:

- YOLO provided better detection for multiple and overlapping objects in the frame.
- SSD showed its strengths in detecting larger objects with precise bounding boxes.
- Some objects were not detected by either model, possibly due to training dataset limitations or insufficient confidence thresholds.

Link to Github Code: Click here.

Answer 3

Solution Description

The task involves identifying whether a given image represents the flag of Indonesia or Poland using basic image segmentation and analysis techniques. Both flags consist of two horizontal stripes, with Indonesia having a red top stripe and Poland having a white top stripe.

I used **YOLOv8n**, a lightweight model from **Ultralytics**, to locate the flag regions within the images. Once the flag region was identified, I applied three different image processing methods to distinguish between the two flags:

- **Image Segmentation and Thresholding**: Dividing the image into top and bottom halves and using a threshold to identify color differences (red vs. white).
- **Sobel Edge Detection and Thresholding**: Applying Sobel edge detection to the grayscale image to find the edges and then comparing the edge strength in the top and bottom halves.
- **Aspect Ratio Method**: Maintaining a 2:3 aspect ratio of the flag (height = 2/3 * width) and analyzing the top and bottom halves for color differences.

While working with these techniques, I encountered several challenges:

- **Model Detection Limitations**: The YOLOv8n model struggled to detect flags in images where there were no other objects or clear edges around the flag.
- **Segmentation Method Limitations**: One image was not identified correctly by the segmentation and aspect ratio methods, but it was successfully identified using Sobel edge detection with thresholding.
- **False Positives with Sobel**: Some images were misidentified using Sobel edge detection, which led to incorrect flag identification.

After experimenting with these methods multiple times, I found that the combination of **image segmentation with thresholding** provided the most reliable results for identifying the flags. Therefore, this approach was ultimately preferred for flag detection.

Explanation of the Code

The code provided contains several key functions used to detect and identify the flags of Indonesia and Poland:

- **'locate_flag_region_with_yolo'**: This function uses the YOLOv8n model to detect objects in the image and identifies the flag region by extracting the bounding box coordinates of the largest detected object. It crops the flag region from the image and returns it.

- **'identify_flag_with_sobel'**: This method applies Sobel edge detection to the cropped flag region. It first converts the image to grayscale, then calculates the Sobel edge magnitude in both the x and y directions. It compares the edge strength in the top and bottom halves of the flag to determine if the top or bottom half contains more edges, which helps in distinguishing between the flags. If the top half contains more edges, it identifies the flag as **Indonesia** (red top), and if the bottom half has more edges, it identifies it as **Poland** (white top).
- **'locate_flag_with_aspect_ratio'**: This function crops the flag region to maintain a 2:3 aspect ratio. It resizes the flag to a standard size and ensures that the flag's height is 2/3 of its width. This method further refines the flag identification by adjusting the aspect ratio.
- **'identify_flag_with_aspect_ratio'**: This method integrates the aspect ratio adjustment with color analysis. After cropping the flag to the correct aspect ratio, it splits the image into the top and bottom halves. It then calculates the mean red intensity in both halves and determines the flag based on the color difference: a stronger red intensity in the top half identifies Indonesia, while a stronger red intensity in the bottom half identifies Poland.
- **'show_images_in_grid'**: This function displays multiple images in a grid for visual analysis. It shows both the original image with the detected flag region highlighted and the cropped flag region in grayscale. It uses the **'identify_flag_with_aspect_ratio'** method to identify the flag in each image.

Key Points to Understand

- **YOLOv8n Model**: The model detects and locates the flag in the image by predicting bounding boxes around the objects. The bounding box with the largest area is assumed to be the flag.
- **Image Segmentation**: Dividing the flag into top and bottom halves to analyze color or edge features.
- **Sobel Edge Detection**: This technique helps to highlight the edges in the image, making it easier to distinguish between flags based on edge strength.
- **Aspect Ratio**: Maintaining the 2:3 aspect ratio of flags helps refine the region for analysis, ensuring that the flag's proportions remain consistent for further comparison.

Challenges and Findings

- **Edge Case Problems**: The YOLO model performed well for images with distinct objects around the flag but struggled with plain flags. In such cases, more advanced edge detection methods like Sobel helped.

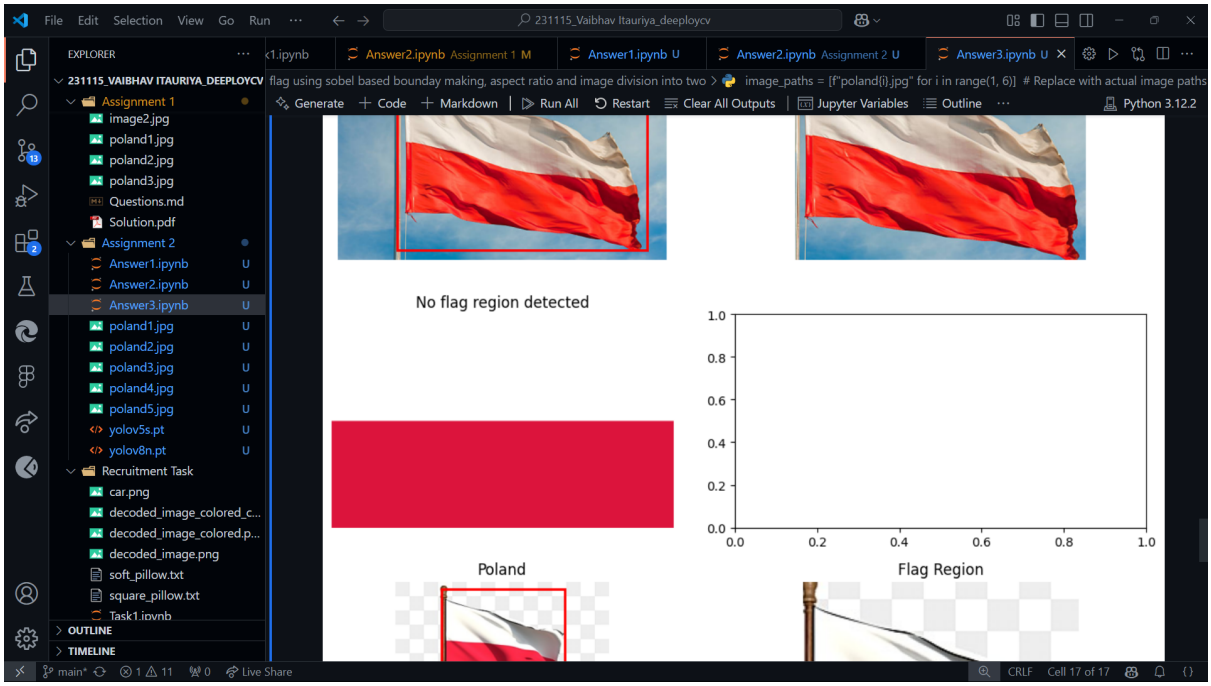

Figure 3: Edge Case Problem

- **False Positives**: Sobel edge detection occasionally produced false positives, misidentifying flags. Image segmentation with thresholding proved to be the most accurate method.

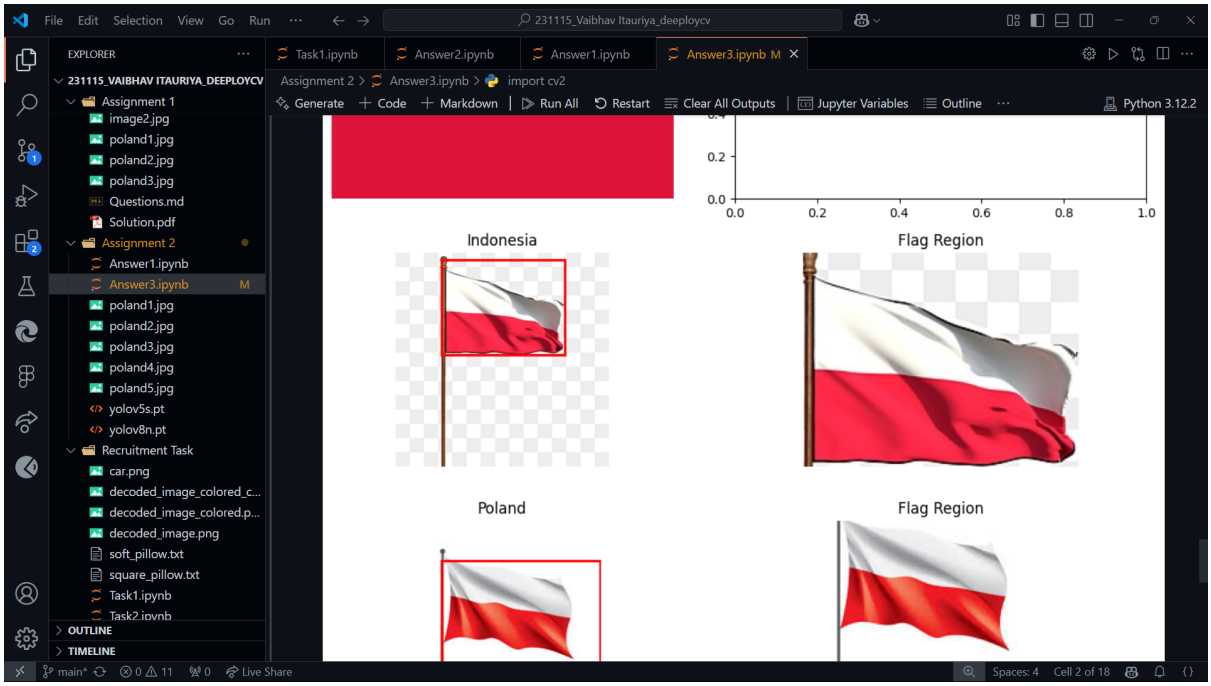

Figure 4: False Positives

Link to Github Code: Click here.