Assignment 2: Prompt Object Segmentation and Tracking

Fariha Rahman (1918608)

Abstract

This report presents two different approaches for video object segmentation and background blurring: (1) Segment Anything Model (SAM) and (2) YOLOv8-Seg. One video was processed using SAM and the second video using YOLOv8. The goal was to isolate the main foreground object while blurring the background in every frame.

1   Introduction

Video object segmentation is widely used in filmmaking, augmented reality, and visual effects. In this assignment, I tested both **SAM** (a prompt-based segmentation model) and **YOLOv8-Seg** (a real-time segmentation model). Each model was applied to a different video to compare segmentation accuracy, speed, and background-blurring performance.

2   Methodology

2.1   1. SAM Segmentation (Video 1)

- Model used: `vit_b`

- A bounding box was manually selected for the target object.

- SAM predictor generated a mask for each frame.

- Morphological closing and dilation were applied to clean the mask.

- Gaussian blur was applied only to regions outside the mask.

- Frames were saved and combined back into an output video.

2.2   2. YOLOv8-Seg (Video 2)

- Model used: `yolov8s-seg.pt`

- YOLO automatically detected and segmented the main object in every frame.

- Polygon segmentation masks were converted to binary masks.

- Background regions were blurred using Gaussian blur.

- The processed frames were exported as a video.

3 Results

3.1 SAM Output

The SAM-based segmentation showed very accurate boundary detection and worked well for frames with complex shapes.

Blurred Background - Frame 0

3.2 YOLOv8 Output

YOLOv8 achieved fast, automatic segmentation across the video. The model performed well even under motion.

4   Comparison

Feature	SAM	YOLOv8-Seg
Initialization	Requires manual bounding box	Fully automatic
Speed	Slower	Faster
Mask Quality	Excellent	Good
Tracking	Manual reuse of prompt	Automatic
Best Scenario	Single clear object	Fast-moving or multi-object scenes

5   Conclusion

SAM provides highly accurate segmentation suitable for detailed video editing, while YOLOv8 offers fast, automated segmentation ideal for dynamic scenes. Both models were able to isolate the foreground object and blur the background successfully, fulfilling all assignment requirements.

Output Videos

- `blurred_video1_SAM.mp4`

- `blurred_video2_YOLO.mp4`