# **YOLOv7 and SAM Background Removal Documentation**

This documentation provides a detailed overview of using the YOLOv7 pose estimation framework in combination with the SAM (Segment Anything Model) to perform background removal on images and videos. This process involves identifying and isolating the foreground (typically human figures) from the background, enhancing the clarity and focus on the primary subjects in media files. Do take note that this only works when there is only one person in the image/video.

Link to Notebook: <https://github.com/Aadharsh1/ML-Deep-Learning/blob/main/Pose_Estimation_Experiments/Yolov7_Experiments/segmentation_pose_estimation.ipynb>

## **Setup and Installation**

**Clone and Setup Repository**:  
  
!git clone https://github.com/RizwanMunawar/yolov7-pose-estimation.git

cd yolov7-pose-estimation

!pip install --upgrade pip

!pip install -r requirements.txt

!wget https://github.com/WongKinYiu/yolov7/releases/download/v0.1/yolov7-w6-pose.pt

**Install Additional Libraries**:  
  
!pip install ultralytics

!pip install 'git+https://github.com/facebookresearch/segment-anything.git'

**Download SAM Model Weights**:  
  
!wget https://dl.fbaipublicfiles.com/segment\_anything/sam\_vit\_h\_4b8939.pth

## **Usage**

### **Remember to change the scripts of pose\_estimate.py whether you want to run this for images or videos. I have included the scripts to be changed in this following notebook.**

### [**https://github.com/Aadharsh1/ML-Deep-Learning/blob/main/Pose\_Estimation\_Experiments/Yolov7\_Experiments/pose\_estimation.ipynb**](https://github.com/Aadharsh1/ML-Deep-Learning/blob/main/Pose_Estimation_Experiments/Yolov7_Experiments/pose_estimation.ipynb)

### **Background Removal for Images**

* **Setup**: Initialize YOLO and SAM models.

**Execution**: Load the image, detect objects using YOLO, and segment the detected object(s) using SAM to create a mask. The background is then removed based on this mask, and the resultant image is saved.  
To remove the background from an image:  
  
image\_name = 'example.jpg' # Replace with your image's filename

image\_path = f'/content/{image\_name}'

img\_name = image\_name.split('.')[0]

output\_path = f'/content/{img\_name}\_bg\_removed.png'

remove\_background(image\_path, output\_path, model, sam, predictor)

### **Background Removal for Videos**

* **Frame Extraction**: Extract frames from the video.
* **Background Removal per Frame**: Apply YOLO and SAM to each frame to remove the background.

**Recompile Video**: Combine the processed frames back into a video.  
To remove the background from a video:  
  
vid\_name = 'example.mp4' # Replace with your video's filename

video\_path = f'/content/{vid\_name}'

original\_frames\_dir = '/content/original\_frames'

processed\_frames\_dir = '/content/processed\_frames'

output\_video\_path = f'/content/{vid\_name.split('.')[0]}\_bg\_removed.mp4'

frame\_count, frame\_rate = extract\_frames(video\_path, original\_frames\_dir)

process\_frames\_with\_progress\_bar(frame\_paths, processed\_frames\_dir, yolo\_model, sam\_predictor)

create\_video\_from\_frames(processed\_frames\_dir, output\_video\_path, frame\_rate)

### **Notes**

* **Performance**: Given the complex nature of video processing and the limitations of the SAM model, processing videos can be notably slow. There is a better way to do this using SAM2, and I have provided the code for that in the same repository.
* **Quality**: The accuracy of background removal depends significantly on the quality of the input video and the model's ability to detect and segment objects accurately.