**Final Project Report: Instance Segmentation and Object Tracking with YOLOv5 and SORT on the DAVIS Dataset**

**1. Project Overview and Objectives**

The aim of this project was to perform instance segmentation and integrate object tracking on video sequences using the YOLOv5 model for detection and the SORT (Simple Online and Realtime Tracking) algorithm for object tracking. Instance segmentation identifies and labels each object instance within frames, while SORT allows tracking the objects across consecutive frames, assigning consistent IDs to each object.

The project objectives were:

* To apply instance segmentation using YOLOv5 on the video frames.
* To implement the SORT algorithm for object tracking, providing consistent object IDs across video frames.
* To evaluate the effectiveness of YOLOv5 and SORT in segmenting and tracking objects within the video sequences.

**2. Brief Description of Methodology**

Brief Description of Methodology: The project utilized the YOLOv5 model for instance segmentation, leveraging the DAVIS dataset, which consists of annotated frames. Key steps included downloading the dataset, converting the annotations into a YOLOv5-compatible format, and processing the data using the YOLOv5 model. Additionally, data augmentation and preprocessing techniques were applied to improve model performance. After YOLOv5 detected and segmented objects in each frame, SORT was applied to link the bounding boxes across consecutive frames. The SORT algorithm uses a combination of the Kalman Filter and Hungarian Algorithm to predict the positions of objects in subsequent frames and match them with new detections.This method allowed for tracking moving objects across the video sequence, assigning unique IDs that remain consistent throughout the video.

**3. Key Results and Achievements**

The key achievements of the project are:

* **Instance Segmentation**:  
  YOLOv5 successfully segmented objects within video frames by using bounding box predictions along with polygon annotations from the dataset. Each object instance in the video frame was accurately segmented.
* **Object Tracking with SORT**:  
  SORT was successfully implemented to track objects across video frames, maintaining consistent object IDs. This allowed the project to extend beyond static segmentation and apply real-time tracking to moving objects, demonstrating the integration of detection and tracking.

**4. Main Challenges Faced and Solutions Implemented**

1. **Converting image annotations into YOLOv5 format.**
   * **Solution**: Custom scripts were developed to convert the dataset’s annotations to YOLOv5 segmentation format, ensuring compatibility with the model..
2. **Managing large datasets efficiently.**
   * **Solution**: Utilized Google Coolab’s resources for data processing and ensured efficient handling of data using optimized pipeline methods.
3. **Tracking Objects Consistently**.  
   Tracking objects across video frames poses challenges, particularly when objects leave the frame or reappear. Matching detections across frames required a robust tracking algorithm.
   * **Solution**: SORT was integrated for real-time tracking. SORT uses a Kalman Filter to predict object positions and the Hungarian Algorithm to associate objects across frames, ensuring smooth and consistent tracking.

**5. Conclusion and Lessons Learned**

This project highlighted the importance of efficient data preprocessing when working with instance segmentation tasks. A key lesson learned was the necessity of adapting datasets to the specific requirements of deep learning models like YOLOv5. Despite the challenges, the project achieved its objectives and laid the groundwork for further enhancements, such as real-time video processing.