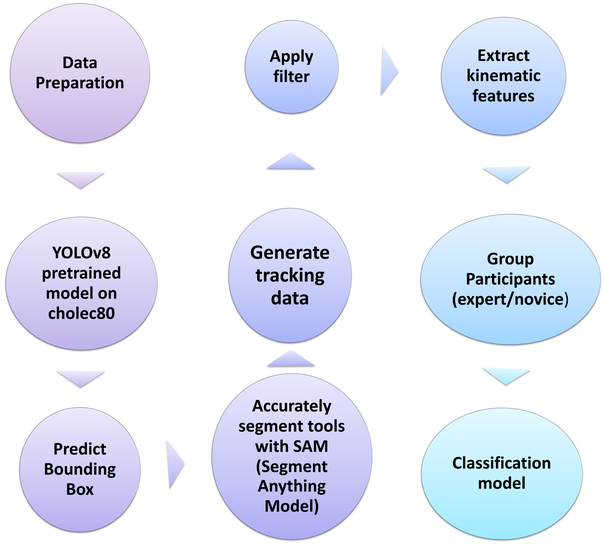
One Working Model is All you Need to Implement Surgical Tool Segmentation and Tracking on Any Procedural Video using Segment Anything Model (SAM)

Shekhar Madhav Khairnar MS, Alexis Desir MD, Carla Holcomb MD, Daniel J Scott MD, Ganesh Sankaranarayanan PhD
Department of Surgery, UT Southwestern Medical Center, Dallas, TX

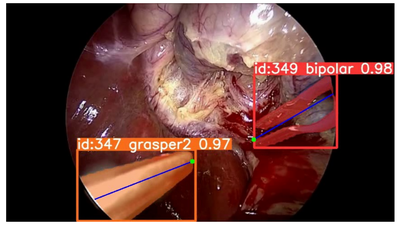
Introduction

- Tool segmentation and tracking is important for automated analysis of surgical procedures from videos for activity detection and assessment.
- Manual annotation is time consuming and not scalable.
- Recent advances in segmentation is the release of the Segment Anything Model (SAM) by Meta AI which is a prompt-driven. segmentation system with zero-shot generalization to unfamiliar objects and images, without the need for additional training.
- The objective of this work is to explore the use of SAM in segmenting surgical tools in laparoscopic video for performance assessment with a case study on simulation videos collected for Nissen fundoplication procedure.

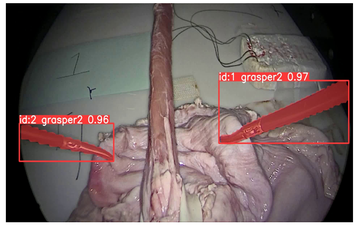
Methods and Procedures



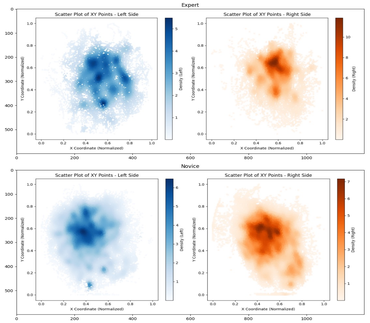
Results



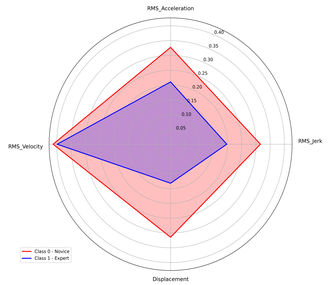
YOLOv8 model trained on Cholec80



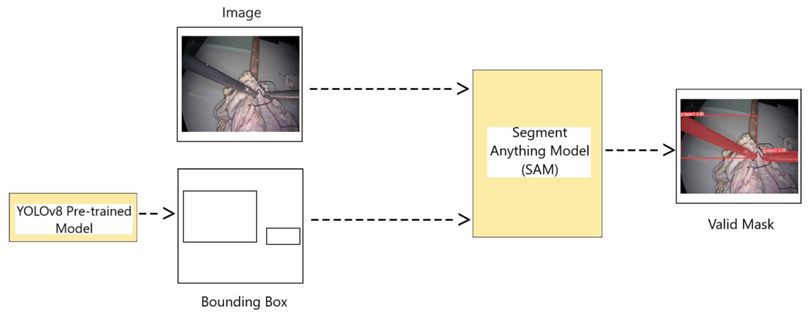
Tool tracking with pretrained model



Expert vs Novice Tool Motion Heatmap



Relevance of Various Kinematic Features



Process Flow

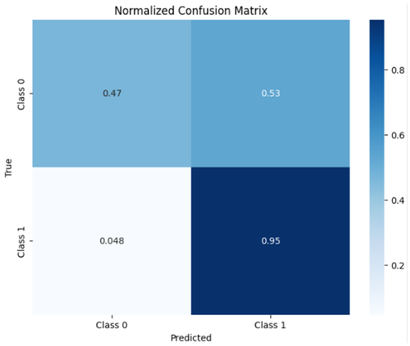
Results

Segmentation:

- F-1 score of 0.99 was achieved at a threshold value of 0.43.
- mAP50-95: The mean Average Precision (mAP) for the range of intersection over union (IoU) thresholds from 50% to 95% was 0.952.

Classification:

- From a total of 38 participants, data for 28 that didn't include any camera movements were analyzed, comprising of 9 novices and 19 experts.
- The analysis utilized a multiple Machine Learning models out of which Logistic Regression performed best with an accuracy of 82% and an F-1 score of 0.8.



Classification Model Confusion Matrix

- From the confusion matrix, we can see that the model performed better for the expert class.
- Among the kinematic features (RMS jerk, RMS acceleration, RMS velocity and displacement), displacement (total path length of the tool) was the most relevant for classification.

Conclusions

- The main purpose of the project is to integrate SAM with a pretrained model for auto annotation.
- We achieved desired performance and was able to segment tools and predict performance.
- With more data, the performance of our model can be improved further.

We acknowledge the support of this from NIH/NIBIB R01 EB025247.