A Robust Multitask Model for Surgical Skill Evaluation

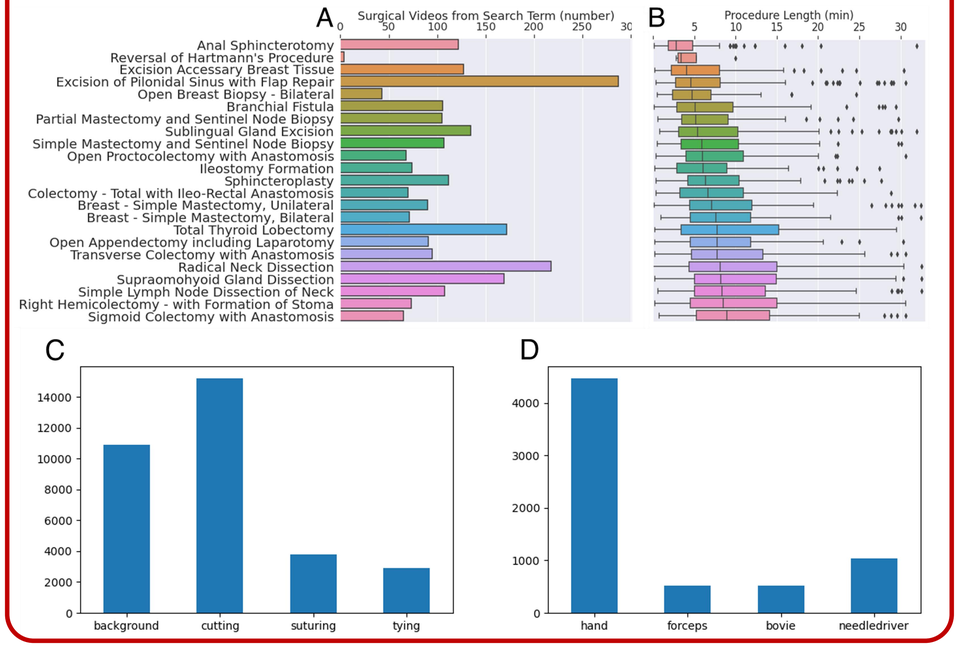
Orr Zohar
Department of Electrical Engineering, Stanford University

Medical AI and ComputeR Vision Lab

Background/Introduction

- Surgical complications are the third leading cause of death globally.
- Open procedures are the most common.
- Can we use videos of surgical videos to evaluate surgical skills?

Dataset

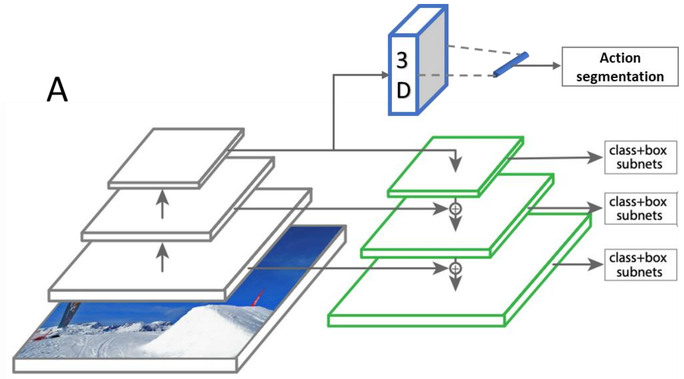


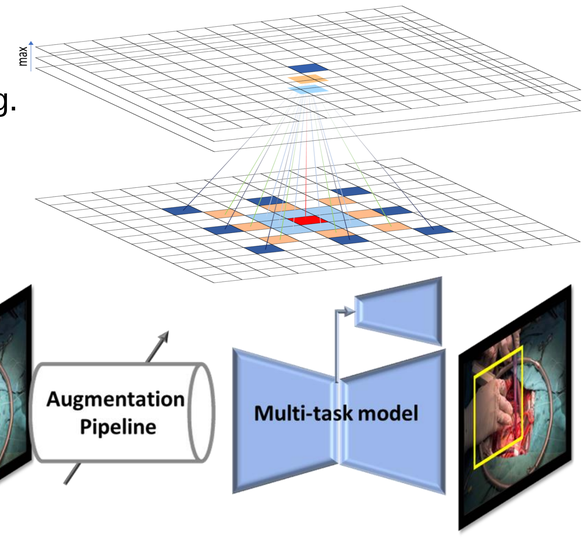
Conclusion & Future Steps

- Multiscale convolution didn't improve scale invariance.
- Robust retraining improved model generalization and performance under domain shift.
- Ongoing clinical trials.

Methods

- Multitask model
- Augmentation pipeline for robustness evaluation & retraining.
- Multiscale convolution for scale invariance.





Experimental Results

- Effect of model robust retraining
- Qualitative results
- Results on new clinical dataset

mAP	bovie	forceps	n. driver	hand	overall
Original	-	0.11	0.06	0.65	0.085
Robust	-	0.25	0.30	0.88	0.285