Creation Of A Computer Vision Model For Anatomic Recognition In Robotic Assisted Minimally Invasive Esophagectomy

Austin Howell MD, Sara Razzaq MD, Maya Sharma, Krithika Mood, Omer Mescioglu, Karishma Muthukumar, Lana Schumacher MD

Tufts University School of Medicine

Boston Medical Center Brighton

Introduction

Despite advances in technology and techniques for esophagectomy, the procedure continues to carry significant morbidity and mortality. Computer vision models assisting trainee surgeons in the recognition of critical anatomy have the potential to reduce associated complications. Our lab has worked to design a process for creation of a computer vision model capable of recognition of anatomy pertinent for robotic assisted minimally invasive esophagectomy (RAMIE).

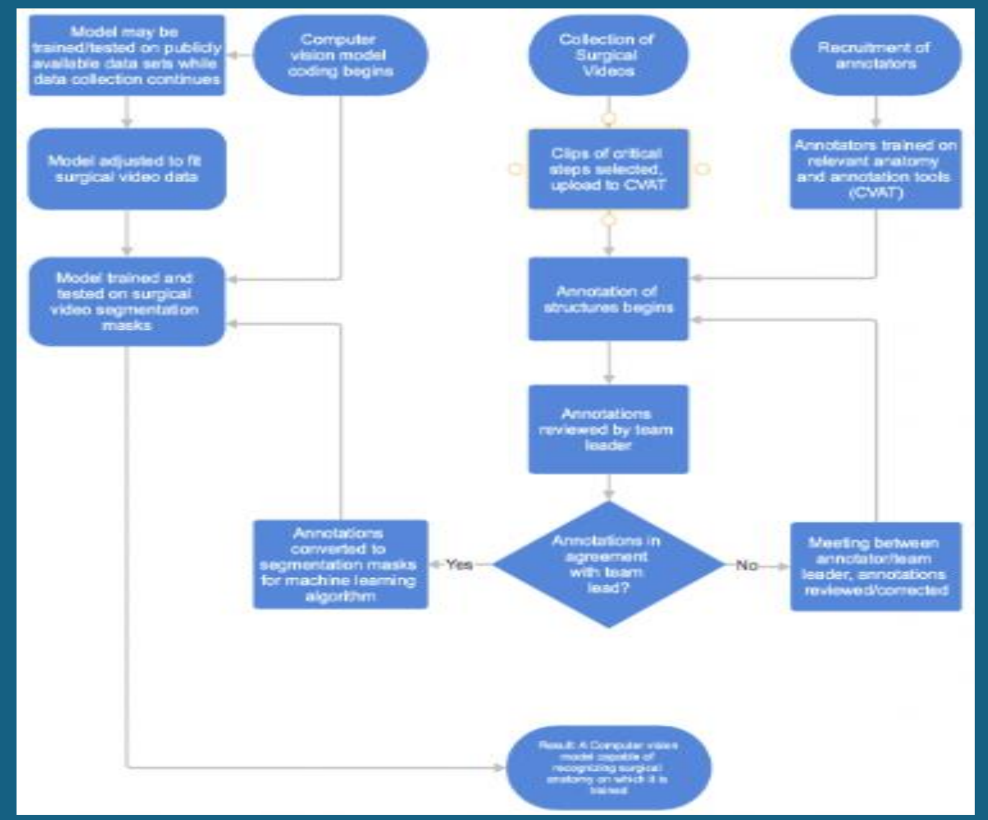

Figure 1: Workflow diagram detailing process for data collection and creation of a computer vision model in RAMIE

Methods

RAMIE videos were recorded on the DaVinci Surgical System and clips of critical surgical steps of the thoracic portion of the procedure uploaded to the computer vision annotation tool (CVAT) for annotation. A team of undergraduate and graduate students underwent training in relevant anatomy and annotation technique by a surgical resident prior to annotation and annotations were reviewed on a weekly basis for accuracy. Currently ten structures(see Table 1) are being evaluated by the model, the surgical videos were also annotated for indocyanine green fluorescence where applicable. Using a U-Net architecture, a computer vision model was designed and tested on a large publicly available dataset for viability as well as our growing dataset as annotation continues.

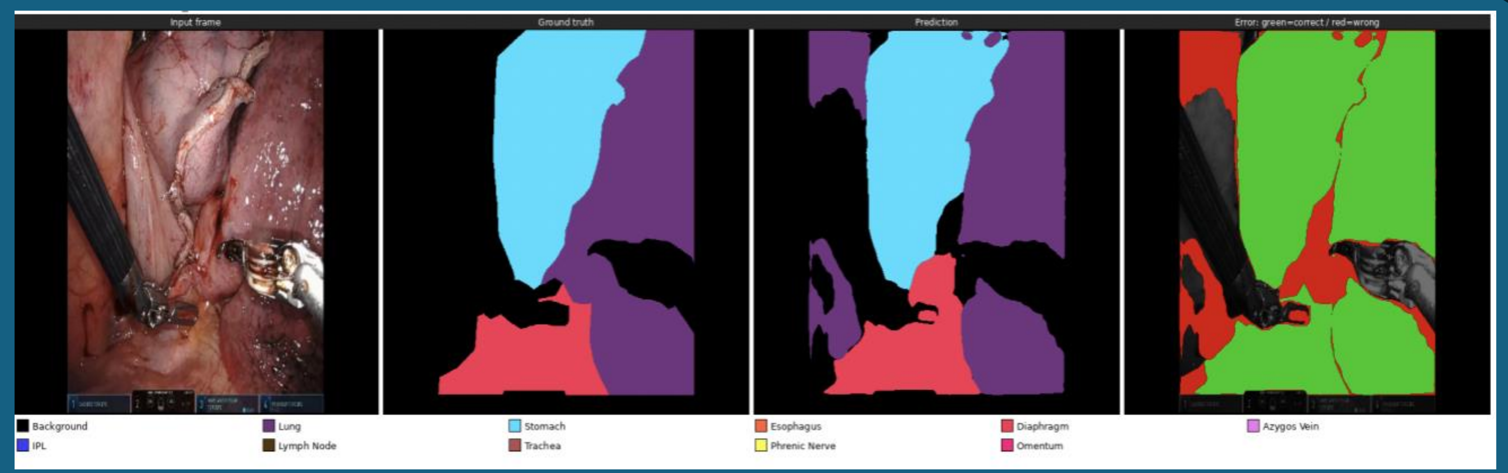

Figure 2: From surgical image to segmentation mask; Images are first uploaded to CVAT, then annotated for key structures before being color mapped onto representational masks. These are used for training a CV model that can make inferences about similar structures

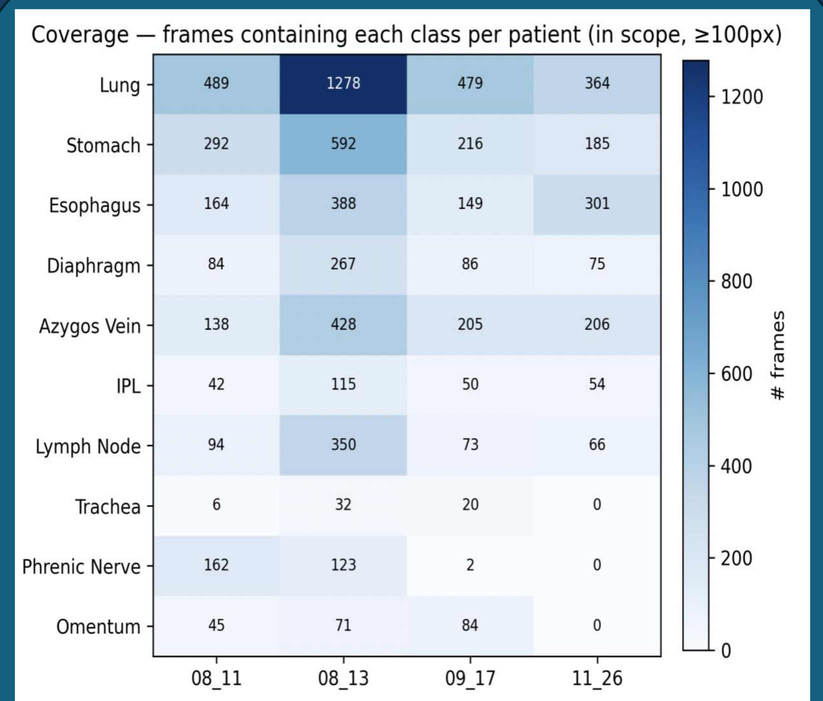

Figure 3: Breakdown of individual class weights by pixel. Common classes such as lung and stomach/gastric conduit greatly outweigh less common ones such as trachea in the training data creating training challenges

Results

The designed model was able to achieve an F1 score averaging 70.09% across three label classes on a publicly available pre-annotated dataset. Preliminary testing on the growing surgical dataset (~2000 frames) shows extremely high variability in the F1 scores across ten structures, each structure ranging from 0.003-0.79 with steady improvement throughout training, further input from

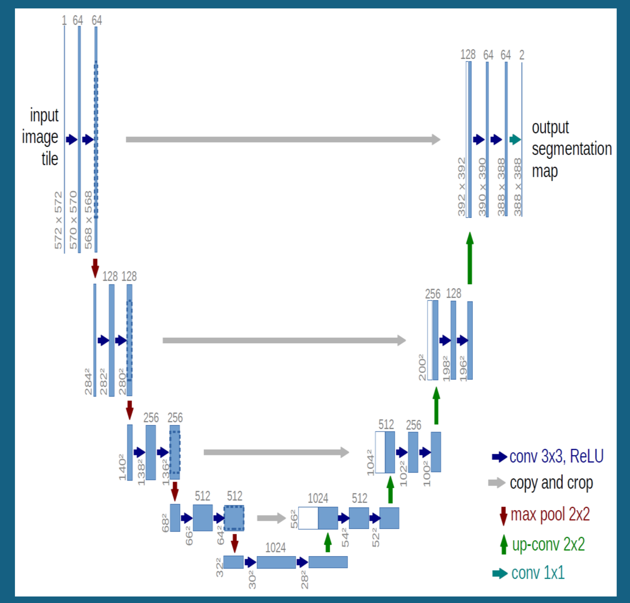

Figure 4: General architecture of a U-net, the neural network being employed by this project

Class	F1 (mean +/- std)	IoU
Lung	0.787 +/- 0.064	0.654
Diaphragm	0.683 +/- 0.142	0.535
Stomach	0.602 +/- 0.022	0.431
Lymph Node	0.516 +/- 0.049	0.349
Esophagus	0.375 +/- 0.049	0.232
Azygos Vein	0.340 +/- 0.133	0.212
IPL	0.323 +/- 0.057	0.194
Omentum	0.277 +/- 0.091	0.164
Phrenic Nerve	0.127 +/- 0.112	0.072
Trachea	0.003 +/- 0.004	0.001

Table 1: F1 Index for tracked classes. Mean F1 correlates strongly with representation within the dataset

Conclusions

This protocol outlines a streamlined process detailing creation of a computer vision model able to be set up with minimal training on a lightweight computer system. The results obtained on existing datasets suggest viability of the model, and the high variability of the results obtained from testing on our own data is not unexpected at this point given the small sample size in certain classes and large number of structures requiring tracking. This is expected to continue to improve greatly as more data becomes available. While our team is focused on RAMIE, the process is generalizable to other computer vision tasks in the field of surgery and represents an accessible starting point to labs interested in exploring computer vision.

Future Aims:

1. Incorporation of AI assisted annotation to significantly speed data collection
2. Improvement of model on both currently studied structures and incorporation of additional structures into inference via increased data and expert computer scientist review
3. Incorporation of ICG fluorescence tracking to assess anastomotic perfusion