Name(s)

Amy Jin

Project Number

S1513

Project Title

Deep Learning-Based Automated Tool Detection and Analysis of Surgical Videos to Assess Operative Skill

Abstract

Objectives/Goals

Annually, 7 million patients suffer surgical complications, many of which are linked to inadequate surgical training and a lack of individualized feedback on how to improve operative technique. To improve surgical quality, it is essential to assess operative skill, a manual process that is time consuming, subjective, and requires experts. Real time automated surgical video analysis would provide a way to objectively and efficiently assess surgical performance. Thus, we aimed to create a deep learning model to perform surgical tool detection, which would enable us to automatically track and analyze tool movements, giving rich insights into operative skill.

Methods/Materials

Since tool localization had not previously been done before, we first had to assemble our own dataset for this task. We designed a MATLAB annotation interface and used it to label 2200 video frames from 15 cholecystectomy surgical videos with the coordinates of spatial bounding boxes around tools. This dataset was then used to train a deep learning model to perform automated tool detection and localization. We leveraged region-based convolutional neural networks (R-CNNs). Input video frames were passed through the VGG-16 convolutional neural network (CNN) and then through a region proposal network (RPN), and the model outputted the spatial coordinates of bounding boxes around surgical instruments.

Results

In comparison with state-of-the-art approaches for automated tool detection, we outperformed existing methods by 23%, improving mean average precision (mAP) from 63.7 to 78.2 using just 10% the amount of training data. Additionally, the network's processing speed at deployment is 5 fps, achieving real time surgical tool detection. We further demonstrated the ability of our method to assess surgical quality by extracting and analyzing key metrics that reflect surgical skill level. In particular, we examined tool usage patterns, tool usage times, movement range, economy of motion, and path length, and used these measures to evaluate surgical performance.

Conclusions/Discussion

To the best of our knowledge, this study was the first to develop an approach for automated surgical video analysis. We collected a new dataset with the spatial bounds of tools, leveraged region-based convolutional neural networks for real time surgical tool detection and localization, and automatically extracted key metrics to assess operative skill.

Summary Statement

We developed a unique approach for automated surgical video analysis, creating a deep learning model to automatically detect and localize surgical instruments and automatically extracting key metrics to assess surgical skill.

Help Received

Serena Yeung, a PhD student at the Stanford Artificial Intelligence Laboratory, answered my questions on how to create the deep learning model, and Dr. Jeffrey Jopling, a surgeon at Stanford Healthcare, provided insight into the metrics that would be most meaningful to analyze.