Deep Learning Object Detection Techniques with Fluoroscopy

IE: 6380 - Deep Learning for Engineering Applications

Josh Friede, John Golec, & Dominik Mattioli
The University of Iowa
Iowa City, IA 52242
josh-friede@uiowa.edu, john-golec@uiowa.edu, & dominik-mattioli@uiowa.edu

March 26, 2019

1 Background, Motivation, & Objective

A surgery simulator project is underway at The University of Iowa, where a team of researchers have focused on simulating the navigation of a 'wire' (i.e. drill) through soft tissue and bone. This simulator currently accommodates two prominent surgeries defined by Wire Navigation: The Dynamic Hip Screw for femoral head fractures (DHS), and The Pediatric Supracondylar Humerus Fracture (PSHF). This simulator is one example of modern research efforts aimed at improving surgical residents' skills before entering the Operating Room – a high-risk and high-cost (emotional, physical, temporal, monetary) environment. The motivation for surgical simulators is to change the traditional paradigm of "on-the-fly" training experience by enhancing skills in a simulated setting, thereby reducing risks and costs.

A general rule-of-thumb with the aforementioned surgeries is to minimize three things: drilling, duration, and radiation (fluoroscopy). The commonly-used evaluation mechanism for surgeries is the Objective Structured Assessment of Technical Skills (OSATS), which, in practice, can be neither objective nor unambiguous. Beyond this, other metrics of "success" are murkier and in need of development. This absence along with a lack of availability of simulator-generated image data are two major research questions in the field which this project is intended to address.

Both the DHS and PSHF surgeries involve tracking the progression of the surgeon's wire throughout the series of fluoroscopic images. This project will focus on the detection of this wire via deep learning object detection techniques. With this capability, image data from novices can be more thoroughly compared and scored against that of experts.

2 The Data

Currently, this project comprises 16 DHS surgeries' and 15 PSHF surgeries' image data generated through a simulator, each composed of 20-100 images (it varies because of each surgeon's preference) for a total of 3,252 images. These images have been analyzed manually, with the coordinates of the wire's tip and point of entry stored in a text file under JavaScript Object Notation (JSON) formatting. From these coordinates, an image mask encompassing the wire via offsets of +/- 10 pixels in the x and y directions is produced. Some images are "negatives", with no wire present.

3 Proposed Method(s)

The team plan to adapt an existing implementation of the Mask R-CNN model, such as Matterport's open-source model built on Python3, Keras, and Tensorflow based on Feature Pyramid Network (FPN) and a ResNet101 backbone. Several of the pre-trained weights from the COCO dataset will be froze in order to decrease the number of trainable parameters. This will allow the training speed for the Mask R-CNN ResNet101 model to increase significantly. To quantitatively evaluate the model results, the team will use the Intersection over Union (IoU) metric, selecting a threshold appropriate for the data and objective.