Surgical tool tracking with object detection for performance assessment in central venous catheterization

Olivia O'Driscoll[1], Rebecca Hisey[1], Daenis Camire[2], Jason Erb[2], Daniel Howes[2],
Gabor Fichtinger[1], Tamas Ungi[1]
[1]Laboratory for Percutaneous Surgery, School of Computing, Queen's University, Kingston, Canada
[2] Department of Critical Care Medicine, Queen's University, Kingston, Canada

Introduction: Medical schools are adopting competency-based medical education, which requires continuous expert supervision to monitor trainee progress. Thus, they seek methods of objectively assessing trainee performance automatically, without expert supervision. Holden et al. revealed that each tool's path lengths are valuable trainee performance metrics in ultrasound-guided interventions, like central venous catheterization (CVC) [1]. We sought to develop a webcam-based approach for performance assessment to eliminate the need for costly systems and expert supervision. In this study, we evaluate the efficacy of using an object detection network for performance assessment in CVC. This project's code can be found here: https://github.com/SlicerIGT/aigt.

Methods: We used the object detection network Faster Region-Based Convolutional Neural Network (Faster R-CNN) to compute the usage time and two-dimensional (2D) path length of seven surgical tools used in central venous catheterization (CVC), as shown in Figure 1. The 2D path length is defined as the sum of Euclidean distances between the centers of sequential bounding boxes for a given tool. Usage time is defined as the number of frames in which a given tool appears divided by the frame rate. Mean average precision (mAP), defined as the area under the precision/recall curve, measured the network's ability to locate and recognize each tool.

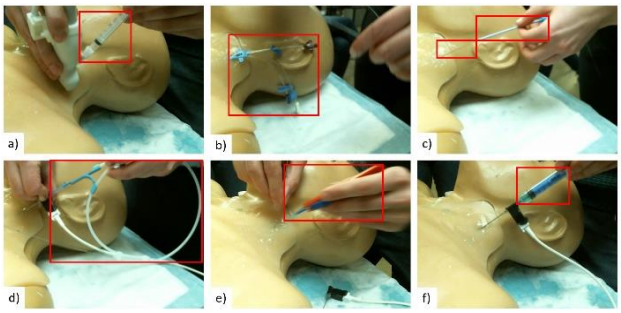

Figure 1. Class ground truth labels: a) anesthetic b) catheter c) dilator and guidewire d) guidewire casing e) scalpel f) syringe

To train the Faster R-CNN, we recorded four medical students each performing five CVC trials on a venous access phantom using a static recording setup. We used a leave-one-user out cross-validation scheme in which the network was trained four times, reserving all five recordings from a single trainee for testing. On each test set, we computed the 2D tool path lengths and usage times for each of the five trials. Comparing the percentage differences in these performance metrics across trials allowed us to evaluate Faster R-CNN's Effectiveness as a performance assessment method in CVC.

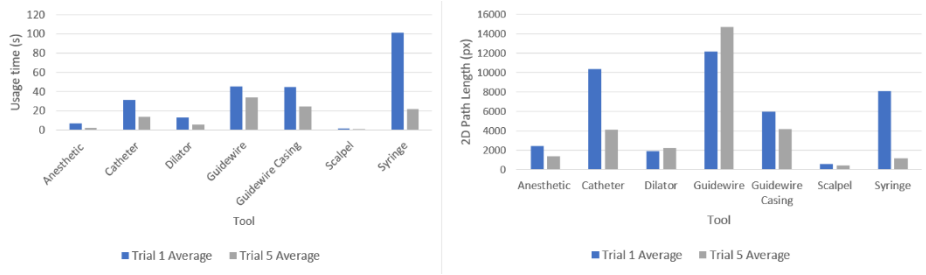

Figure 2. Comparisons of a) Tool usage times and b) 2D tool path lengths in trials 1 and 5

Results: The average tool usage time reduction in the students' first and fifth procedural trials was 52%, and that of 2D path length was 29%, as shown in Figure 2. The neural network's mAP was 0.66, with tool-specific mAPs between 0.49 and 0.92.

Conclusions: Faster R-CNN recognized procedural tool accurately enough to compute metrics that demonstrated improvement with the trainee's skill level. Faster R-CNN is an effective skill-assessment method in CVC.

Acknowledgements: This work was funded in part by NIH/NIBIB and NIH/NIGMS (via grant 1R01EB021396-01A1), by CANARIE's Research Software Program, and is supported as a Collaborative Health Research Project (CHRP #127797) by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Institutes of Health Research (CIHR). R. Hisey is supported by the Queen Elizabeth II Graduate Scholarships in Science and Technology (QEII-GSST). G. Fichtinger is supported as a Canada Research Chair.

References:

[1] Holden, M., Xia, S., Lia, H., Keri, Z., Bell, C., Patterson, L., Ungi, T. and Fichtinger, G. "Machine learning methods for automated technical skills assessment with instructional feedback in ultrasound-guided interventions," International journal of computer assisted radiology and surgery, 1993-2003 (2019)