Ultrasound Probe Pose Classification for Task Recognition in Central Venous Catheterization

Colton Barr, Rebecca Hisey, Tamas Ungi, Gabor Fichtinger

Laboratory for Percutaneous Surgery, School of Computing, Queen's University, Kingston, Canada

INTRODUCTION: Due to the restricted time and availability of expert observers, medical trainees are often limited in the feedback they receive when practicing a new skill. Central Line Tutor is a training system for central venous catheterization (CVC) that aims to alleviate this problem by automatically providing instruction and feedback to students using computer vision-based task recognition [1]. As students practice CVC on a simulated patient, the system monitors their progression through the workflow using a combination of an RGB camera and electromagnetic (EM) tracker. While almost all tasks can be detected in RGB images based on the presence of a particular tool, distinguishing between different ultrasound tasks requires knowledge about the pose of the ultrasound probe. Using an EM tracker to extract ultrasound (US) probe pose, however, effectively doubles the overall cost of the system and reduces the feasibility of deploying Central Line Tutor for widespread use. In order to remove the need for EM tracking and substantially reduce the overall cost of the Central Line Tutor system, we propose the use of deep learning to classify US probe orientation in real-time from streamed US images.

METHODS: A classification architecture was created that combined a U-Net with a CNN-based classifier. Our dataset was generated by segmenting the carotid and jugular vessels in 20 tracked US sequences collected from 4 trainees. Each sequence yielded approximately 1100 US images for a total dataset of 22000 images. The images were also labelled with the corresponding US probe orientation relative to the major vessels of the neck. The probe poses were categorized as either "long-axis", "cross-section" or "undefined". Both the U-Net and subsequent pose classifier were tested using a leave-one-user-out scheme, where one trainee's data was reserved for testing each round and the rest of the data was used for training and validation. For each fold, the U-Net and classifier were trained on the same dataset, followed by testing using the same test set. This ensured that both networks were tested on novel data in each fold. The U-Net performance was quantified using Intersection over Union (IoU), which measured the overlap between the human segmentations and network predictions, while classification performance was quantified using accuracy, precision and recall.

RESULTS: The mean testing set IoU score for U-Net cross-validation was 0.727 ± 0.045. The mean classification performance on the testing set across 4 folds was an accuracy of 0.905 ± 0.036, with a weighted mean precision of 0.913 ± 0.008 and weighted mean recall of 0.905 ± 0.038 across all 3 classes. Figure 1 illustrates two sets of US images captured with different probe poses, as well as the U-Net performance on these images.

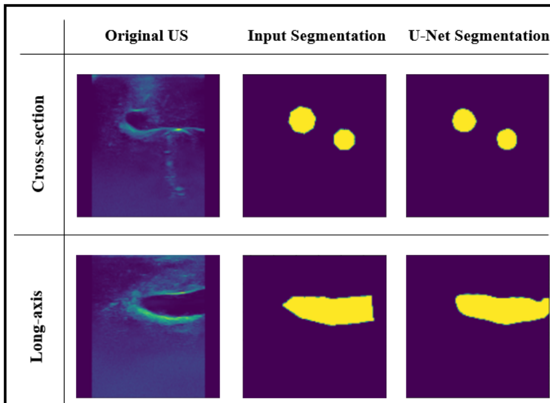

Fig. 1. Demonstration of Cross-section and Long-axis probe poses with their corresponding US image, ground truth segmentation and U-Net segmentation.

CONCLUSIONS: The classification of US probe pose directly from US images can be achieved with a high level of accuracy using a combined U-Net and CNN classification model. This model's performance demonstrates that US pose classification may be suitable for recognizing US tasks during central line insertion, and future work will be aimed at testing the network's performance in real-time task recognition within Central Line Tutor. All code associated with this project is available at github.com/SlicerIGT/aigt .

ACKNOWLEDGEMENTS: C. Barr was supported by an NSERC CGS M award. R. Hisey is supported by the QEII-GSST award. G. Fichtinger is supported as a Canada Research Chair. Work funded, in part, by NIH/NIBIB and NIH/NIGMS (via grant 1R01EB021396-01A1 - Slicer+PLUS: Point-of-Care Ultrasound) and by CANARIE's Research Software Program. This work also was financially supported as a Collaborative Health Research Project (CHRP #127797).

REFERENCES:

[1] Hisey, R., Ungi, T., Holden, M., Baum, Z., Keri, Z., McCallum, C., Howes, D. and Fichtinger, G. "Real-time workflow detection using webcam video for providing real-time feedback in central venous catheterization training," Proc. SPIE 10576, Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling, 1057620 (2018)