

Reinforcement learning approach for video-based task recognition in central venous catheterization

Rebecca Hisey¹, Brian Chen², Tamas Ungi¹, Daenis Camire³, Jason Erb³, Daniel Howes³, Gabor Fichtinger¹

¹Laboratory for Percutaneous Surgery, School of Computing, Queen’s University, Kingston, Canada

²Medical Informatics Laboratory, School of Computing, Queen’s University, Kingston, Canada

³Department of Critical Care Medicine, Queen’s University, Kingston, Canada

Introduction: Central venous catheterization (CVC) involves inserting a catheter into a major vein such as the internal jugular. This essential skill is taught during residency for many medical specialties such as critical care. The workflow for this procedure is long. By recognizing the task that is being done, we can provide feedback about workflow compliance that will help novices learn the procedure. Here, we present a method for recognizing CVC tasks from a video using a combination of convolutional neural networks (CNN) and reinforcement learning.

Methods: In our initial attempts to recognize the tasks in CVC, we recognize tasks solely based on the tool that is identified using a CNN with a soft-max output layer. For this approach, we use the output of the same CNN and use it to train a policy for predicting the most likely task. We model this problem as a search for the optimal path through a grid. An optimal path is the path that yields the highest reward. Each row of the grid represents one of the eight tasks that we are recognizing, and each column represents a single frame of the video (Fig. 1). To generate a policy, we begin by using the CNN to classify all frames in 4 videos of medical residents performing CVC. In addition to classifying all frames with a CNN, we also manually label each frame based on the task that is being done. Since each tool corresponds one-to-one with a task in the procedure, we define our reward scheme by using the CNN’s confidence that each tool is being used and add an additional bonus of +2 for the tool that corresponds to the task label (Fig. 2). Next, for each video in the training set, we use Q-learning to train a policy to find the optimal path based on our reward scheme. We allow the system to train for 100 000 episodes, or until it converges to the optimal reward.

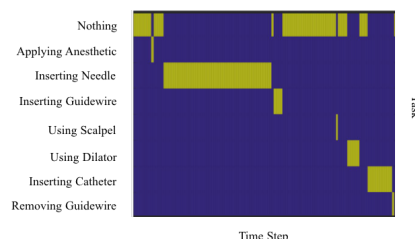


Figure 1. Example of an optimal path for a single video.

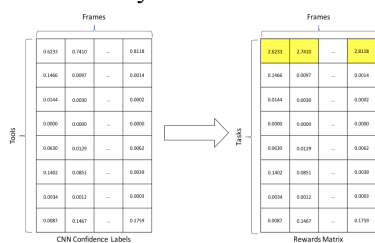


Figure 2. Generation of reward matrix.

To use the policy for prediction, we classify each frame in the test video and find its nearest neighbor in the training set. The nearest neighbor is found by comparing the output of the CNN along with the time at which it occurs in the video. The system then uses the policy that was generated in training to select the task that has the highest expected reward. For this work, we perform a leave-one-video-out cross validation whereby we use 4 videos to train the policy, and test on the remaining video. This process is repeated for all 5 videos and we report the average accuracy.

Results: The average accuracy across all cross-validation folds was 85%. As can be seen from the example Figure 3, the path found by our system closely resembles the optimal path shown in Figure 1. In contrast, the path found using the CNN alone, is much more prone to errors and only achieves 56% accuracy on average.

Conclusions: Our approach using CNNs and reinforcement learning clearly outperforms the CNN alone. This is due to the fact that the reinforcement learning approach takes temporal information into account. This allows it to still correctly predict the task even when the CNN recognizes the wrong tool based on which task is likely being performed at that time in the procedure. Further work will be done to evaluate the applicability of this work to real-time task detection.

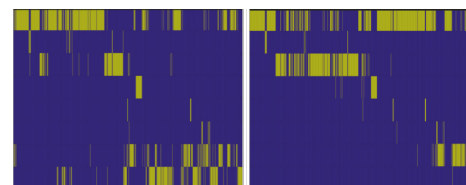


Figure 3. Predicted Paths using CNN alone (Left) and CNN & Q-Learning (Right).

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 Ontario Graduate Scholarship. G. Fichtinger is supported as a Canada Research Chair.