ACS 2023 Surgeons and Engineers: A Dialogue on Surgical Simulation Meeting

Research Abstracts

Developing Artificial Intelligence Models for Medical Student Suturing and Knot-tying Video-Based Assessment and Coaching

Ganesh Sankaranarayanan, PhD; Babak Namazi, PhD; Madhuri Nagaraj, MD, MS, and Daniel J. Scott, MD FACS

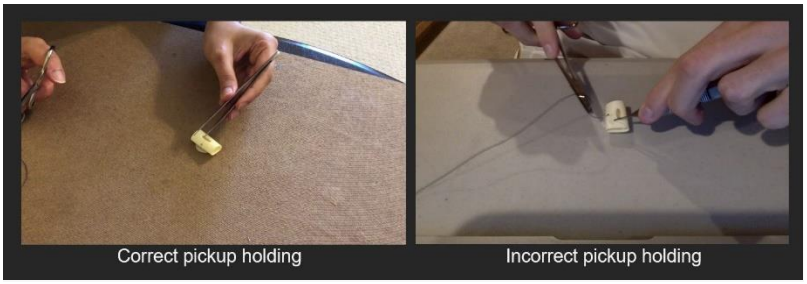
UT Southwestern Medical Center, Dallas, Texas

Introduction: During COVID-19, the in-person suturing curriculum for our second-year medical students was adapted to an at-home video-based assessment model, which required the need for an automated assessment and feedback tool. We aimed to use artificial intelligence to develop an assessment as well as error based formative feedback tools for our curriculum.

Methods: Second-year medical students were asked to submit a video of a simple interrupted knot on a penrose drain with instrument tying technique after self-training to proficiency (performing the task under two minutes with no critical errors). Videos were manually annotated with a pass-fail grade and 9 steps of the procedure. We developed and trained two deep learning models to identify improper instrument hold and improper knot-tying. For instrument holding error, 16 frames were randomly selected from the "Needle loading and insertion" step of each video to train a Convolutional Neural Network (CNN). To identify knot-tying errors, a sequence of 64 consecutive frames (approximately 1 frame per second) from the knot-tying steps were used to train a 3D CNN. The inputs to our 3D CNN model were the calculated optical flow frames of each video.

Results: A total of 229 medical student videos were manually reviewed (150 pass, 79 fail). Of those who failed, critical error distribution was 15 knot-tying errors, 47 instrument-holding errors (Fig. 1) and 17 both errors. Task performance time in median [interquartile range] was significantly shorter for passing group at 53 seconds [43.25-64.75] as compared to the failing group at 62 seconds [51-76.5] (p-value <.01). A total of 216 videos were used to train the CNNs. A k-fold cross validation (k=10) was used for training and validation. The accuracy of the instrument holding error detection model is 88% with an F-1 score of 73%. For the knot-tying model, the accuracy is 89% with an F-1 score of 50%.

Conclusions: Using deep learning, we have developed a system for assessment and directed feedback to better acquire open surgical suturing skill.



Correct pickup holding

Incorrect pickup holding