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

Research Abstracts

Video-based-assessment of Surgical Skill Using Unsupervised Temporal Attention

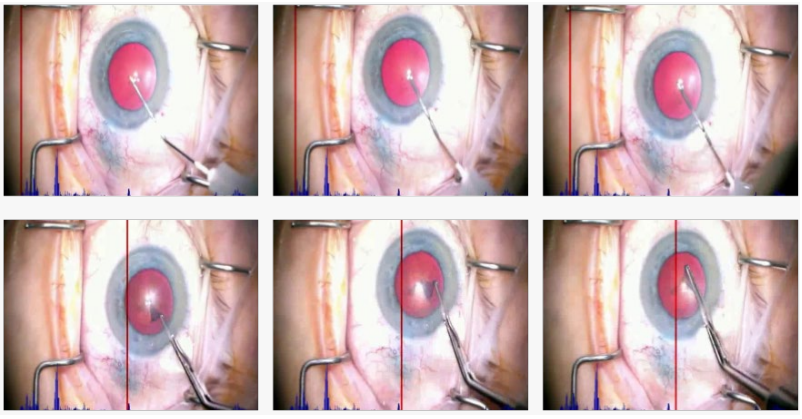
Bohua Wan, Meng; Michael Peven, PhD; Gregory Hager, PhD; Shameema Sikder, MD, and S. Swaroop Vedula, MBBS, PhD, MPH

Johns Hopkins University School of Medicine, Baltimore, MD; Johns Hopkins Medicine Wilmer Eye Institue, Baltimore, MD

Introduction: Algorithms for video based assessment (VBA) of surgical skill should be explainable to facilitate their adoption. Our objective was to develop and validate deep learning algorithms for VBA of surgical skill that localize predictive segments of surgical activities.

Methods: We used a dataset of 99 videos of capsulorhexis for which ground truth (expert/novice) was specified by an expert surgeon using the International Council of Ophthalmology's Ophthalmology Surgical Competency Assessment Rubric for phacoemulsification. We first trained a CNN as a feature extractor from video images. We then stored the features to save memory, which would allow us to compute frame-level features for videos of any length (limited only by storage capacity). We then trained a simple long short-term memory network with temporal attention modules using the extracted frame-level features as input. The temporal attention facilitates localization of segments in the video that influenced the predicted skill. We used 5-fold cross-validation to estimate algorithm performance in terms of sensitivity and specificity, and qualitatively analyzed influential video segments identified by the network.

Results: Our network had a sensitivity of 0.75 (95% confidence interval [CI], 0.61 to 0.84) and a specificity of 0.71 (95% CI, 0.57 to 0.82). The figure shows a visualization of the computed temporal attention. The red vertical line denotes the time position of the current frame. The blue bars at the bottom of the video show the importance of the corresponding frames.



Conclusions: Temporal attention, normalized across the entire video, is beneficial for both performance and explainability of deep learning methods for VBA of surgical skill.