ACS 2025 Surgeons and Engineers: A Dialogue on Surgical Simulation Meeting

P-E-06

Research Abstracts

Comparison of YOLOv5 and YOLOv8 in Detection and Classification of Needle and Needle Driver States in Suturing Training
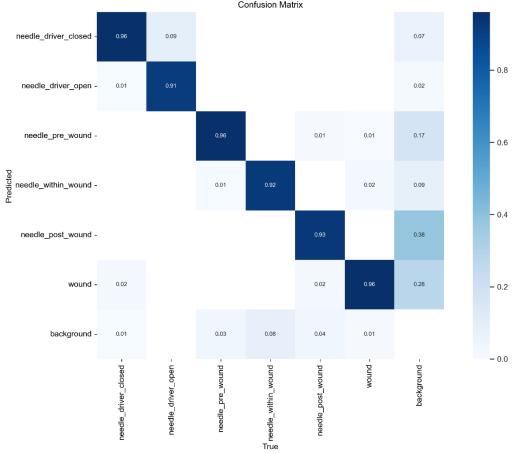
Stevan Craig Fairburn

University of Alabama Birmingham, Birmingham, AL

Introduction: Accurate detection and classification of surgical instruments are critical for enhancing surgical education and improving training outcomes. This study aims to develop a tool that provides quantitative feedback to students, allowing them to improve their suturing skills and techniques. By leveraging advanced object detection models like YOLOv5 and YOLOv8, the project enhances surgical education through precise and automated instrument state recognition.

Methods: A dataset of 777 images from suturing videos on synthetic pads was annotated, capturing six distinct classes: "needle_driver_closed," "needle_driver_open," "needle_pre_wound," "needle_within_wound," "needle_post_wound," and "wound." The dataset was divided into five folds for cross-validation. YOLOv5 and YOLOv8 models were trained on each fold with performance metrics such as precision, recall, mean Average Precision (mAP) at IoU threshold 0.5 (mAP@50), and mAP across IoU thresholds from 0.5 to 0.95 (mAP@50-95) calculated to assess the models' performances.

Results: The YOLOv5 model achieved an average precision of 0.9536, recall of 0.9549, mAP@50 of 0.9699, and mAP@50-95 of 0.6572. The YOLOv8 model, tested for comparison using the same parameters, achieved a mAP@50 of 0.973, mAP@50-95 of 0.668, precision of 0.934, and recall of 0.941. These outcomes demonstrated reproducibility and consistency across the YOLO models.

Conclusions: Both YOLOv5 and YOLOv8 models provide robust detection and classification of needle and needle driver states in suturing training. The consistent outcomes across these models underscore their potential for integration into surgical education programs. Future steps include hyperparameter optimization, video segment isolation, and implementing a user-friendly tool for surgical trainees.

Confusion Matrix