Automated Surgical Skill Assessment Using Tooltip Motion Metrics

Farzad Aghazadeh¹, Bin Zheng², Mahdi Tavakoli³, Hossein Rouhani¹
¹Department of Mechanical Engineering, University of Alberta, Edmonton, Canada
²Department of Surgery, University of Alberta, Edmonton, Canada
³Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada

*hrrouhani@ualberta.ca

ABSTRACT

Surgical complications can jeopardize patients’ safety and impose significant costs on healthcare systems. One potential solution to reduce surgical complications is to ensure that surgeons and surgical residents acquire the required technical skills before performing surgeries, which can be fulfilled by surgical skill assessment. Surgical skill assessment is primarily conducted by senior surgeons who observe surgical trainees’ performance during or after the operations. However, this approach is labour-intensive, subjective, and prone to bias. To address this shortcoming, we aimed to objectively assess surgical skills based on tracked motion data of surgical instruments.

In this study, we recruited seven participants without any prior surgical experience and four expert surgeons with extensive surgical experience. Participants performed peg transfer, bimanual peg transfer, and rubber band translocation tasks. Surgical graspers’ tooltip path length, path length along the instrument axis, mean linear speed, mean linear acceleration, and motion smoothness (derived from dimensionless tooltip motion jerk) were calculated. These five tooltip motion metrics were normalized to be in the range of 0 to 1, and the dimensionality of the motion metrics was reduced using principal component analysis (PCA). Afterwards, tooltip motion data were inputed to three algorithms, i.e., k-nearest neighbours (KNN) with k = 5, logistic regression (LR), and support vector machine (SVM) to classify experts and novices.

Our results showed that KNN, LR, and SVM had classification accuracy (TP+TN / TP+TN+FP+FN) values of 0.64, 0.91, and 1, respectively. Furthermore, KNN, LR, and SVM had F1-score (2*TP / 2*TP+FP+FN) values of 0.5, 0.89, and 1, respectively. Therefore, SVM represented the highest accuracy and F1-score for the expert surgeon classification. It should be noted that since tooltip motion metrics significantly differed between expert surgeons and novices, SVM could achieve accuracy and F1-score values of 1. However, the inclusion of surgical residents, who presumably have intermediate skill levels, in the classification algorithms can reduce the accuracy and F1-score values.

In conclusion, using surgical instrument motion metrics along with the SVM algorithm, accurate surgical skill classification was obtained. In future studies, tooltip motion metrics of surgical residents will be added to the classification algorithms. This is because expert surgeons- and novices-surgical residents classification is more challenging than experts-novices classification. Moreover, eye-tracking data should be included in future studies as eye-hand coordination plays an essential role in surgical proficiency.