Multimedia Tools and Applications (2020) 79:30111–30133
https://doi.org/10.1007/s11042-020-09540-y

Proposing novel methods for gynecologic surgical action recognition on laparoscopic videos

Toktam Khatibi[1] · Parastoo Dezyani[1]

Received: 8 February 2020 / Revised: 15 July 2020 / Accepted: 4 August 2020 /
Published online: 13 August 2020

Abstract

Laparoscopy or minimally-invasive surgery (MIS) is performed by inserting a camera called endoscope inside the body to display the surgical actions online with the ability to record and archive the video. Recognizing the surgical actions automatically from the laparoscopic videos have many applications such as surgical skill assessment, teaching purposes, and workflow recognition but is a challenging task. The main aim of this study is proposing novel automatic methods for surgical action recognition from the laparoscopic video frames. For this purpose, three different scenarios are designed, evaluated and compared using 5-fold cross validation strategy. The first and the second scenarios are based on deep neural networks and combination of pre-trained CNNs and conventional machine learning models, respectively. The last scenario combines handcraft feature extraction, pre-trained CNNs, feature engineering based on complex networks and conventional classifiers. Dataset analyzed in this study is ITEC LapGyn4 Gynecologic Laparoscopy Image dataset. Experimental results show that the second and the third scenarios have highly desirable performance for multi-instance surgical action recognition with the average accuracy of 99.20 and AUC of 99.12. On the other hand, for single-instance surgical action recognition, the third scenario outperforms the compared ones with the average accuracy of 99.05 and AUC of 96.41. Moreover, different feature sets in the third scenario are ranked and assigned the importance score based on "Mean Decrease of Accuracy" measure. The first-ranked features are the deep features extracted from our proposed CNNs in the first scenario and the second-ranked ones are the features engineered from the complex networks.

Keywords Minimally-invasive surgery (MIS) · Medical image processing · Multi-instance classification · Deep neural networks · Wrapper feature selection · Feature engineering

✉ Toktam Khatibi
toktamk.khatibi@modares.ac.ir

[1] School of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran