Internship : Automatic Evaluation of a Flexible Endoscopy Training Session

keywords : Deep Learning, Neural Networks, Action Recognition

IRCAD and IHU are two partner institutes aiming at creating new tools for the surgery of the future, notably through a common Research and Development team.

IRCAD (Institut de Recherche contre les Cancers de l'Appareil Digestif) constitutes a reference in the domain of assisted digestive surgery, specifically in the area of minimally invasive surgery, via the use of Augmented Virtuality and Augmented Reality. The IHU (Institut de Chirurgie Guidée par l'Image) in Strasbourg develops an innovative surgery to improve medical care of the patients, with a personnalized approach combining the best minimally invasive technologies with the latest progress in medical imaging. For many projects, the R&D team of the IHU works jointly with the R&D team of IRCAD.

In the last decade, the use of flexible endoscopy has increased significantly, morphing from a diagnostic tool to becoming a new modality for ultra minimally invasive surgery. This procedure requires highly skilled surgeons, and training tools ensure that graduating surgical residents acquire the essential flexible endoscopic skills.

Virtual Reality simulators are used, however they are expensive, not widely available and often lack physical realism. The Basic Endoscopic Skills Training (BEST) Box is a new experimental low-cost solution, developed at the ihu Strasbourg, that is an alternative to simulation. It is a cheap to manufacture and is used to practice and evaluate six surgery tasks (two shown in Figure 1). Furthermore, the program has recently developed to include an evaluation system that generates a user competence rating. However, this needs an expert surgeon in order to rate and identify the accomplished and failed tasks, and cannot be run in real-time.

a) Forward peg transfer   b) Puncturing

Figure 1 : Some tasks of the Basic Endoscopic Skills Training (BEST) Box

The objective of this internship is to develop a deep learning solution to help bridge this gap and allow for automatic real-time performance scoring independent of, and without the need for a trained human rater. This proposed study will involve the development, training and validation of a dedicated neural network which would be able to automatically score the performance of a user on the BEST Box. As architecture, the student will be responsible for experimenting with, and extending a state-of-the-art Video Action Recognition network. Video Action Recognition does not yet have an established optimum solution [1,2,3,4,5], and the student will have to investigate, and develop the right one suited for this application. Some of other works on endoscopy surgery action detection and user evaluation are highly relevant [6,7,8,9,10].

Technical Background: Research Skills, Image Processing, Machine Learning, Statistics and Deep Learning

Coding Skills: python (required), C++ (desired), CUDA

Deep Learning Frameworks: Tensorflow, Pytorch, Keras, Caffe, MXNet

[1] Jeff Donahue et al., Long-term Recurrent Convolutional Networks for Visual Recognition and Description, CVPR 2015.

[2] CDC: Convolutional-De-Convolutional Networks for Precise Temporal Action Localization in Untrimmed Videos - Z. Shou et al, CVPR2017.

[3] Cascaded Boundary Regression for Temporal Action Detection - Jiyang Gao et al., BMVC 2017.

[4] Temporal Action Detection with Structured Segment Networks - Y. Zhao et al., ICCV2017

[5] SST: Single-Stream Temporal Action Proposals - S. Buch et al, CVPR2017.

[6] Zia et al., Temporal clustering of surgical activities in robot-assisted surgery, IJCARS 2017

[7] Fawaz et al., Evaluating surgical skills from kinematic data using convolutional neural networks, MICCAI 2018.

[8] Kassahun et al., Surgical robotics beyond enhanced dexterity instrumentation: a survey of machine learning techniques and their role in intelligent and autonomous surgical actions. International Journal of Computer Assisted Radiology and Surgery 11(4), 553–568 (2016)

[9] G. Yengera, D. Mutter, J. Marescaux, N. Padoy, Less is More: Surgical Phase Recognition with Less Annotations through Self-Supervised Pre-training of CNN-LSTM Networks, arXiv:1805.08569, 2018

[10] A.P. Twinanda, G. Yengera, D. Mutter, J. Marescaux, N. Padoy, RSDNet: Learning to Predict Remaining Surgery Duration from Laparoscopic Videos Without Manual Annotations, *IEEE Transactions on Medical Imaging (TMI)*, to appear, arXiv preprint, 2018

Supervisors :
Güinther Saibro, Ingénieur de Recherche
Toby Collins, Directeur de Recherche
Equipe R&D IRCAD/IHU
1, place de l'hôpital, 67091 Strasbourg
e-mail: guinther.saibro@ihu-strasbourg.eu
Gratification : 1100 € gross/month
Duration : 6 months
Internship start date : 1st quarter of 2019
Workplace : IRCAD – Strasbourg – Centre-ville