ON REDUCING EFFORT IN EVALUATING LAPAROSCOPIC SKILLS

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ
Institute of Information Technology
itec.aau.at

Sabrina Kletz, Klaus Schoeffmann

Institute of Information Technology (ITEC) - Alpen-Adria-Universität (AAU), Klagenfurt, Austria
sabrina@itec.aau.at, ks@itec.aau.at

ABSTRACT

Training and evaluation of laparoscopic skills have become an important aspect of young surgeons' education. The evaluation process is currently performed manually by experienced surgeons through reviewing video recordings of laparoscopic procedures for detecting technical errors using conventional video players and specific pen and paper rating schemes. The problem is, that the manual review process is time-consuming and exhausting, but nevertheless necessary to support young surgeons in their educational training. Motivated by the need to reduce the effort in evaluating laparoscopic skills, we investigate state-of-the-art content analysis approaches for finding error-prone video sections.

PROBLEM STATEMENT

A laparoscopic surgery is a complex sequence of surgical activities and human errors can occur. Such errors do not necessarily have fatal consequences, but repeated mistakes can only be avoided by recognizing them. The most frequent errors in laparoscopic Roux-en-Y gastric bypasses [1] and hysterectomies [2] occur during

- abdominal access of trocars and instruments,
- grasping and dissection,
- suturing,
- and use of energy devices.

Automatic detection of technical errors in videos of laparoscopic procedures is not a trivial task and is accompanied by many research challenges in vision-based object and action recognition:

- Large variations in appearance of objects (e.g., viewpoint, scale, occlusion, orientation, illumination, camera motion)
- Conditions (e.g., specular reflections, blurriness, smoke, blood)
- Strong object motion (i.e., errors occur in a fraction of a second)

OBJECTIVES AND CONTRIBUTION

The purpose is to identify error-prone surgical activities for highlighting relevant sections in videos to reduce the surgeon's effort in evaluating laparoscopic skills. The research question is:

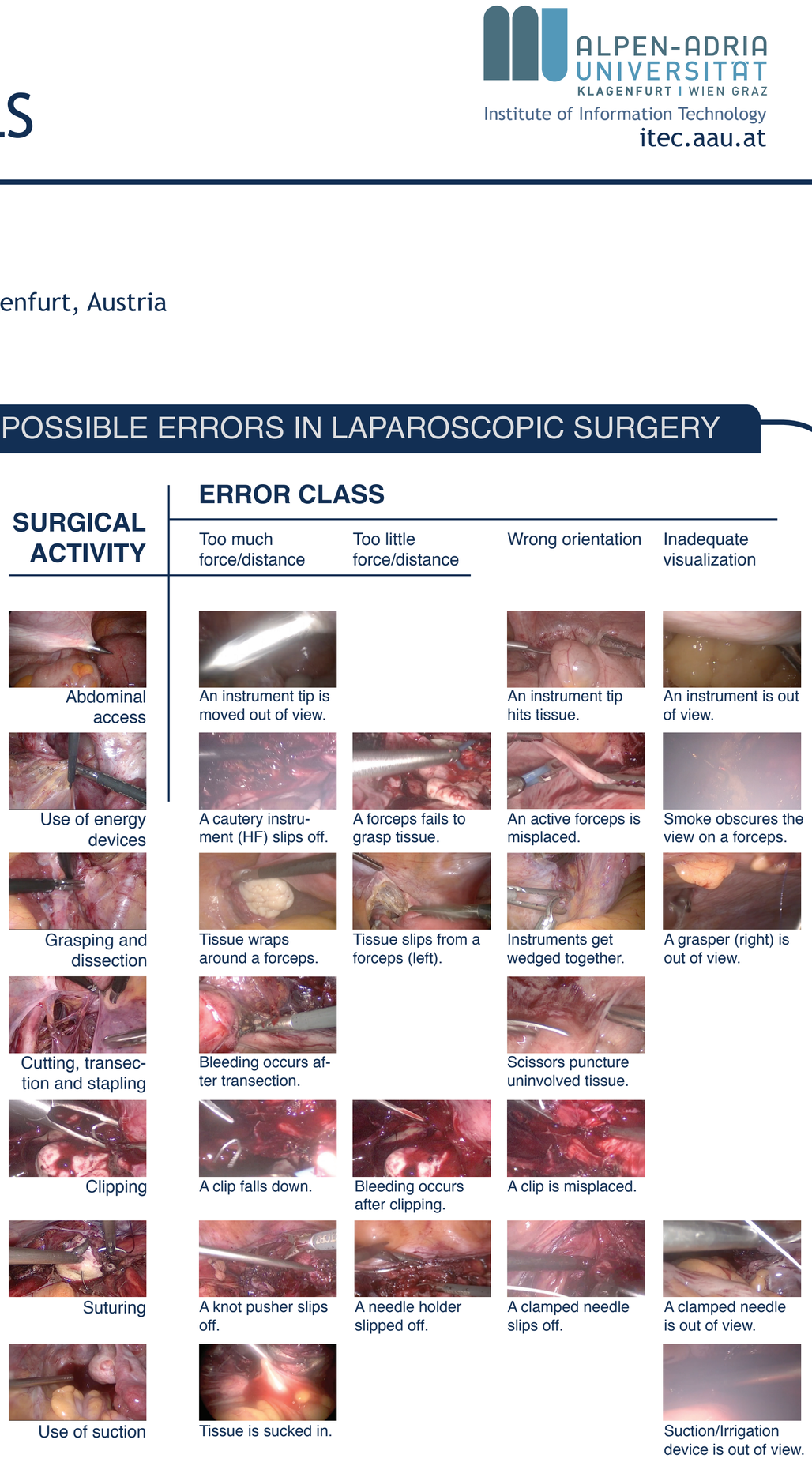
To what extent can visual object detection methods be used to identify error-prone surgical activities with reasonable accuracy?

We **investigate surgical activities and deep learning approaches** and the following research objectives will be addressed by the mentioned research question:

1. Evaluating instrument detection methods with a focus on high precision by investigating and improving upon state-of-the-art convolutional neural networks (CNNs).
2. Designing and developing a technical skill evaluation prototype to integrate applicable object detection approaches.
3. Evaluating whether technical errors are recognizable with reasonable precision by learning spatio-temporal dependencies of moving objects.

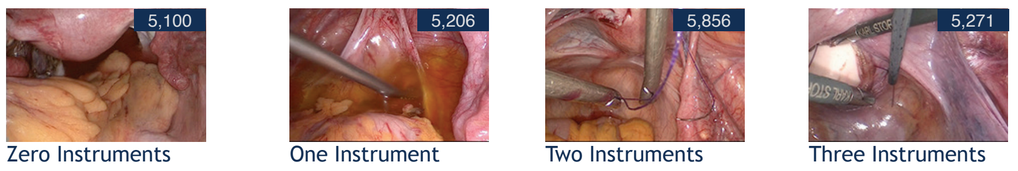
POSSIBLE ERRORS IN LAPAROSCOPIC SURGERY

SURGICAL ACTIVITY	ERROR CLASS: Too much force/distance	Too little force/distance	Wrong orientation	Inadequate visualization
Abdominal access	An instrument tip is moved out of view.		An instrument tip hits tissue.	An instrument is out of view.
Use of energy devices	A cautery instrument (HF) slips off.	A forceps fails to grasp tissue.	An active forceps is misplaced.	Smoke obscures the view on a forceps.
Grasping and dissection	Tissue wraps around a forceps.	Tissue slips from a forceps (left).	Instruments get wedged together.	A grasper (right) is out of view.
Cutting, transection and stapling	Bleeding occurs after transection.		Scissors puncture uninvolved tissue.	
Clipping	A clip falls down.	Bleeding occurs after clipping.	A clip is misplaced.	
Suturing	A knot pusher slips off.	A needle holder slipped off.	A clamped needle slips off.	A clamped needle is out of view.
Use of suction	Tissue is sucked in.			Suction/Irrigation device is out of view.



WORK IN PROGRESS

In [3] we present different datasets addressing specific problems in gynecologic laparoscopy; among them one is for classifying different amounts of visible surgical tools composed of ~22k sample images and categorized into four content classes:



Zero Instruments | One Instrument | Two Instruments | Three Instruments

Results of baseline evaluations show promising accuracy in identifying the number of instruments within images:

GoogLeNet	Jaccard Index	Recall	Precision	Specificity	Accuracy	MCC*	F1-Value
Zero Instruments	0.862	0.928	0.924	0.976	0.965	0.903	0.926
One Instrument	0.673	0.790	0.819	0.944	0.907	0.743	0.804
Two Instruments	0.631	0.763	0.785	0.921	0.878	0.69	0.773
Three Instruments	0.770	0.898	0.844	0.946	0.934	0.827	0.870
Weighted Average	**0.730**	**0.842**	**0.841**	**0.946**	**0.920**	**0.787**	**0.841**

*MCC = Matthews Correlation Coefficient

REFERENCES

[1] Bonrath, E. M., Zevin, B., Dedy, N. J., & Grantcharov, T. P. (2013). Error Rating Tool to Identify and Analyse Technical Errors and Events in Laparoscopic Surgery. British Journal of Surgery, 100(8), 1080–1088.

[2] Husslein, H., Shirreff, L., Shore, E. M., Lefebvre, G. G., & Grantcharov, T. P. (2015). The Generic Error Rating Tool: A Novel Approach to Assessment of Performance and Surgical Education in Gynecologic Laparoscopy. Journal of Surgical Education, 72(6), 1259–1265.

[3] Leibetseder, A. and Petscharnig, S. and Primus, M. J. and Kletz, S. and Münzer, B. and Schoeffmann, K. and Keckstein, J. (2018). LapGyn4: A Dataset for 4 Automatic Content Analysis Problems in the Domain of Laparoscopic Gynecology. Proceedings of the 9th ACM Multimedia Systems Conference, ACM.

Acknowledgments

We thank Dr. Husslein for helpful discussions and insights into this topic. This research is supported by Alpen-Adria-Universität, Klagenfurt and Lakeside Labs GmbH, Klagenfurt, Austria and funding from the European Regional Development Fund and the Carinthian Economic Promotion Fund (KWF) under grant KWF 20214 u. 3520/ 26336/38165.