ÓBUDAI EGYETEM
ÓBUDA UNIVERSITY

THESIS BOOKLET

RENÁTA NAGYNÉ ELEK

Endoscopic image and kinematic data-based autonomous technical and non-technical skill assessment in Robot-Assisted Minimally Invasive Surgery

Robotsebészeti technikai és nem-technikai készségek automatizált mérése endoszkópos kameraképek és kinematikai adatok alapján

Supervisor:
Prof. Dr. Tamás Haidegger

DOCTORAL SCHOOL OF APPLIED INFORMATICS AND APPLIED MATHEMATICS

Budapest, 2023

Contents

Frequently used Notations and Symbols

TABLE 1: Common abbreviations and notations

AI	Artificial Intelligence
AR	Augmented Reality
CG	Control Group
CIS	Computer-Integrated Surgery
ConvAuto	Convolutional Autoencoder
CNN	Convolutional Neural Network
CSRT	Channel and Spatial Reliability Tracker
DCT	Discrete Cosine Transformation
DFT	Discrete Fourier Transformation
DoA	Degree of Autonomy
DoF	Degrees of Freedom
DNN	Deep Neural Network
DVRK	Da Vinci Research Kit
dVSS	da Vinci Surgical System
ECM	Endoscope Camera Manipulator
FCN	Fully Convolutional Neural Network
FDA	United States Food and Drug Administration
FLS	Fundamentals of Laparoscopic Surgery
FPS	Frames per Seconds
FN	False Negative
FP	False Positive
FRS	Fundamentals of Robotic Surgery
GEARS	Global Evaluative Assessment of Robotic Skills
HMM	Hidden Markov Model
HR	Heart Rate
ICARS	Interpersonal and Cognitive Assessment for Robotic Surgery
iDT	Improved Dense Trajectory
IEC	International Electrotechnical Commission
IoU	Intersection over Union
ISO	International Organization for Standardization
JIGSAWS	JHU–ISI Gesture and Skill Assessment Working Set
KT	Knot-tying
LC	Laparoscopic Cholecystectomy
LDA	Linear Discriminant Analysis
LDLJ	Natural log of Dimensionless Jerk
LOOCV	Leave One Out Cross-Validation

LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MES	Medical Electrical System
MIS	Minimally Invasive Surgery
MP	Medical Professionals
MSE	Mean Squared Error
NASA-TLX	NASA Task Load Index
NN	Neural Network
NP	Needle-passing
NTS	Non-Technical Skills
NOTSS	Non-Technical Skills for Surgeons
OF	Optical Flow
OR	Operating Room
OSATS	Objective Structured Assessment of Technical Skill
PCA	Principal Component Analysis
PnP	Perspective n Point Transformation
PSM	Patient Side Manipulator
RAMIS	Robot-Assisted Minimally Invasive Surgery
ResNet	Residual Neural Network
ROI	Region of Interest
ROS	Robot Operating System
SURG-TLX	Surgical Task Load Index
SA	Situation Awareness
SDS	Surgical Data Science
SPARC	Spectral Arc Length
ST	Suturing
STIP	Space Temporal Interest Points
SVM	Support Vector Machines
TCP	Tool Center Point
TN	True Negative
TP	True Positive
VR	Virtual Reality
2/3D	Two/Three Dimensional

1 Background

In the case of Minimally Invasive Surgery (MIS), surgeons reach the human organs through small skin incisions. This approach – compared to open-access surgery – typically results in less tissue trauma, smaller scars and quicker recovery for the patient [1]. On the other hand, MIS requires advanced surgical training, because handling the MIS surgical (laparoscopic) tools is not intuitive, the operating area is visualized by an endoscopic camera, the surgeon has limited range of view, and long surgeries can cause great fatigue. As of today, MIS surgical skill assessment is not part of the clinical training practice. Robot-Assisted Minimally Invasive Surgery (RAMIS) can provide help to these MIS challenges. The da Vinci Surgical System (dVSS) is at the moment the market leading RAMIS system with more than 7500 clinical systems across the world [2]. DVSS can help the surgeons with 3D vision, ergonomy, intuitive tool handling, tremor filtering and motion scaling. DVSS is a teleoperational system, the surgeon remotely operates the patient-side robotic arms from a master console. With dVSS, MIS can be delivered more precisely, and it can decrease the workload on the surgeon. MIS was a revolution in medicine 30 years ago, and it is now part of the everyday clinical routine. During their training, residents have to take practical exams, but the assessment of these is usually manually done by an expert surgeon. There are technical and non-technical skills in surgery; technical skills are such as: instrument handling, knowledge of equipment, knowledge of procedure and indirect indicators (used forces, elapsed time), etc.; non-technical skills are such as leadership, situation awareness, decision making, dealing with stress, etc. (Fig. 1). Both skillsets are required for a good surgeon, and with frequent assessment, personalized training and better patient outcome can be achieved. The goal is to assess the skills of the surgeon autonomously, with available or additional sensors. With autonomous approaches, non-biased, objective surgical skill assessment can be achieved, furthermore, it does not require extensive time commitment from expert surgeons [3]. With the dVSS, these sensory data

Fig. 1. Surgical skills are composed of technical skills, non-technical skills and workload management. Surgical skill training and assessment can directly affect the skills of the surgeon, which derive better patient outcome.

are more straightforward to access, the kinematic and video data of the surgeon can be recorded, and the motion of them can be examined. Kinematic data-based skill classification in RAMIS can be realized with an almost perfect accuracy, but endoscopic images are available in traditional MIS and in training videos, it is beneficial to examine skill assessment through both endoscopic images and kinematic data.

1.1 Skill assessment in RAMIS

The improvements of RAMIS can help the execution of the surgery, however, RAMIS is still a hard task to master; continuous training and feedback about the performance is crucial. Technical skills in RAMIS are related to the basic skills of the surgeon, the control of the robot and MIS tools and tissue handling. Nevertheless, non-technical skill (NTS) assessment is less deterministic. Despite the fact that RAMIS can decrease the mental workload of the surgeon as shown by several studies, RAMIS remains a challenging operation to perform not just physically, but mentally as well, because of the constant communication, teamwork, leadership, decision making and workload conditions [4, 5, 6].

The workload on the surgeon – which represents the effort to perform a task – can be high in several segments of a procedure: there are mental, physical and temporal demands; furthermore, task complexity (including multitasking, task novelty), situational stress and distractions can influence the outcome of the surgery [7, 8]. Naturally, the same task can cause different workload to different operators. NTS is related to the workload which can directly affect surgical outcome. While it is straightforward that technical skills are crucial, NTS can be as important as technical skills [9, 10, 11].

In the literature, three main approaches for surgical performance assessment can be identified [12, 13]:

- self-rating questionnaires;
- expert-based scoring and
- automated skill assessment.

Questionnaires are filled by the operator; thus, it is easy to implement, yet subjective. Objective scoring is done by an expert panel, based on a standardized method [14]. Expert ratings are supposedly objective, yet may be biased for personal reasons; furthermore, they can be hard to implement, being human resource intensive. Automated skill assessment is based on objectively measurable parameters (such as applied forces, movement velocity, etc.). However, in most cases, it is technically not easy to implement. Robotic surgical systems can provide a unique platform for objective skill assessment due to the recordable kinematic and video data [15]. The mentioned surgical skill as-

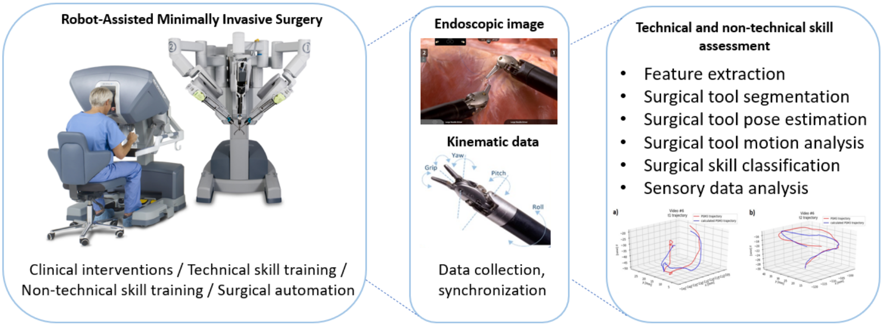

Fig. 2. Autonomous technical and non-technical surgical skill assessment input (endoscopic image and kinematic data) and steps. All steps under skill assessment are covered in this PhD thesis.

sessment approaches can be found in technical and non-technical skill assessment as well. For the latter, questionnaires and automated solutions can be useful tools, and for technical and NTS assessment, all of the methods (questionnaires, expert-rating and automated techniques) can be utilized.

In this PhD thesis, I am focusing on autonomous technical and non-technical surgical skill assessment based on endoscopic image and kinematic data (Fig. 2). Sensory data in RAMIS can originate from clinical interventions, technical or non-technical skill training, or even surgical subtask automation. Under autonomous skill assessment, image feature extraction, surgical tool segmentation, surgical tool pose estimation, motion analysis and skill classification techniques were considered.

2 Research Goals

Training and skill evaluation is crucial in the case of MIS, since it requires extensive practical skills as well, while surgical skill assessment is not yet part of the clinical practice. With autonomous technology-supported approaches, non-biased, objective surgical skill assessment can be achieved, furthermore, it does not require resources from human evaluations.

While kinematic data-based skill classification in RAMIS can attain close to 100 % accuracy, but endoscopic images are more generally available in traditional MIS and in training videos, therefore there is a need for RAMIS technical skill assessment examined through endoscopic images. However, the accuracy of image-based surgical skill assessment is still below the kinematic data-based solutions. Surgical skill assessment has to be validated on an annotated database, but for the widely-used JHU–ISI Gesture and Skill Assessment Working Set (JIGSAWS), image annotation is not available, thus it is hard to validate image-based skill assessment.

- *Problem 1:* Image-based skill assessment's accuracy should be improved, because endoscopic camera images are the only widely accessible data. Semantic segmentation of the surgical tools on training videos should be available.

Non-technical surgical skill assessment is not a widely studied research domain, while clinical failures in the OR may just as commonly originate from low non-technical skills of the surgeon than the lack of technical skills.

- *Problem 2:* Non-technical skill needs training and evaluation. Non-technical surgical skills should be examined with autonomous techniques as well.

Automating the motion of the camera holder arm can decrease the cognitive workload on the surgeon, while surgical skill assessment is essential in this cases for safety reasons. However, image-based skill assessment can be complex if the camera is moving as well.

- *Problem 3:* Image-based skill assessment should be the part of automated camera motion for safety reasons. A solution should be provided to filter the camera motion from the visual scene.

The problems identified above summarize three areas covering an important set of interconnected issues, and address scientific problems relevant to the clinical practice.

3 Materials and Methods

3.1 Image-based automated technical skill assessment in RAMIS

In the case of classical MIS or the widely available training and surgical videos, only 2D endoscopic image is accessible. In my first Thesis group, a surgical tool pose estimation technique was proposed for the dVSS articulated tools, providing algorithmic support for autonomous technical skill assessment (in dissertation: Chapter 3). The main reason for using generated image-based positions is that kinematic data-based solutions typically provide much higher accuracy for skill classification, compared to endoscopic image-based solutions. The idea behind this research is that if accurate tool pose can be generated from the images, it can be a good alternative of kinematic data-based solutions. The tool pose estimation is based on 2D endoscopic images, shape features and iterative Perspective n Point transformation method [16]. The introduced technique was validated on the Synthetic MICCAI dataset, which contains training videos and kinematic data on the dVSS [17]. The smoothness of the generated tool trajectories were examined with the natural log of Dimensionless Jerk (LDLJ) and Spectral Arc Length (SPARC) [18, 19]. My method's usability for skill classification was validated on the JIGSAWS dataset [20] with Random Forest-based time series supervised classifier [21]. JIGSAWS dataset is the most widely used dataset for surgical skill assessment, yet, there has been no solution for semantic segmentation of the surgical instruments in it. In this Thesis, different neural networks based on UNet [22], LinkNet [23] and TernausNet [24] were examined and applied to segment surgical tools in the JIGSAWS dataset aiming to lead

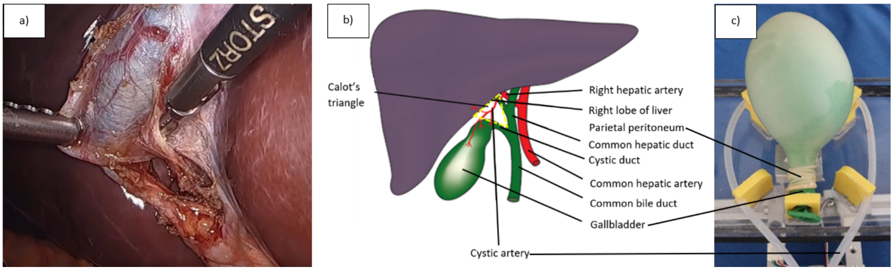

Fig. 3. Laparoscopic cholecystectomy (LC) surgical scene and anatomy. A) LC surgical scene, after exposing the Calot's triangle; b) Anatomy of the cyst and its environment; c) Surgical phantom created for LC with the peritoneum, cyst and the cystic artery.

image-based surgical skill assessment, where the ground truth was generated semi-autonomously with Optical Flow (OF), which tracks the apparent motion of objects in consecutive frames of a video sequence.

3.2 Automated non-technical skill assessment and workload evaluation in MIS

NTS assessment can be as important as technical skill assessment in MIS, while it is not widely studied in the clinical practice, and it is not included routinely in training. In this Thesis group, I proposed a methodology for autonomous NTS assessment and workload evaluation in laparoscopic cholecystectomy (LC) training (Fig. 3), (in dissertation: Chapter 4). An automated, image and force sensor-based NTS skill classification solution is presented, aiming to understand the correlations behind tool motions, used forces and situation awareness, dealing with stress and distractions. For this, a laparoscopic training platform was created, simulating certain parts of LC: clipping the cystic artery and the dissection of the parietal peritoneal layer. The training task requires situation awareness, dealing with stress and distractions, and decision making. 20 trials were recorded, performed by 7 subjects with different surgical experience. The analysis of the workload was done with a self-rating questionnaire (Surgery Task Load Index, SURG-TLX) [25]. Tool motion was processed with an object tracking computer vision algorithm (CSRT [26]). The time series data (force and image) were classified with a Fully Convolutional Neural Network-based (FCN) classifier [27].

3.3 Skill assessment and automation

Since RAMIS typically operates on soft tissues, automation in this domain can be extremely difficult, due to the constantly changing environment. At the moment, it is rare to see RAMIS systems on the market with autonomy over Level of Autonomy [15]. Related research is mainly focusing on surgical subtask automation, which belongs to partial automation, where the automation of surgemes and motion primitives are necessary as well. However, the workflow of RAMIS contains subtask elements, where choosing the proper subtask can be extremely hard, since it can be critical regarding the patient outcome. These subtasks can be monotonous and time-consuming, thus automation of them could decrease the workload on the surgeon as well. Skill assessment in the case of automation can be crucial for safety reasons. In this Thesis group, a complete framework for OF-based surgical skill assessment in the case of autonomous endoscope motion is introduced to assess surgical skills while automation is employed (in dissertation: Chapter 5).

For surgical subtask automation, a marker-based visual servoing method was proposed to automate camera motion in RAMIS, which can help the surgeon focusing only on the instrument control in the da Vinci Research Kit (DVRK) enhanced research mode. The usage of visual markers ensured the elimination of a significant fraction of vision-related errors, thus, during the development and validation, the main focus could remain on the robot control aspect of the visual servoing problem. The proposed method used eye-in-hand visual servoing approach [28]; it allows the autonomous movements of the camera by receiving an expected goal function (expected view). It relies on stereoscopic images as an input, the positioning of the Endoscope Camera Manipulator (ECM) operates in 3D. Such a video-based surgical skill assessment method can be widely used, since training videos are available. In my work, a two dimensional endoscopic image data-based surgical skill assessment method is developed, which can be used on any available surgical training videos, if the videos are annotated, including the references to surgical skills. For skill assessment, OF was chosen as an objective parameter, because of its robustness and reasonable computational performance. If the camera is automated, with proper OF ego-motion filtering (which is introduced in the third Thesis group), the skills of the surgeon can be measured with the proposed method. The OF features were processed for skill classification using feature extraction techniques: CNN [29], Long Short-term Memory (LSTM) [30], ResNet [31], Convolutional Autoencoder (ConvAuto) [32], Discrete Fourier Transformation (DFT) and Discrete Cosine Transformation (DCT) [33]. The results were validated on the JIGSAWS dataset with LOOCV. OF-based skill assessment can be a much more complex problem, when the viewpoint is moving as well; OF techniques cannot make a difference between motion originated from moving objects in the space and from the self-motion of a moving camera. The motion of the viewpoint – ”ego-motion” – has to be extracted from the OF vector field to detect moving objects in the space. The method utilizes the robot’s state of motion and depth information as an input. The motion of the camera is extracted from the OF vector field by re-projecting the 2D pixels to the 3D world, and to extract the re-projected 3D OF with the camera holder arm’s known state of motion.

4 New Scientific Results

Thesis group 1

In my first thesis group, I proposed solutions for robotic surgery technical skill assessment based on endoscopic image data.

Thesis 1/I: I proposed a surgical tool pose estimation technique and semantic segmentation algorithm for the da Vinci Surgical System's articulated tools, applicable to autonomous technical skill assessment. My surgical tool pose estimation method provides a generic solution, since it does not require kinematic data or a complete model of the tool. It performed on 2D endoscopic images, based on shape features and iterative Perspective n Point transformation method. The method's accuracy was proven in surgical skill assessment, where it outperformed the state of the art on knot-tying videos (89.33 % accuracy).

Thesis 1/II: Ground truth generation for semantic segmentation for the JIGSAWS dataset was introduced, which was not available before. Using deep learning methods, I provided an accurate solution for semantic segmentation of the surgical tools for the JIGSAWS dataset, for image-based skill assessment (97.38 % accuracy, 79.91 % dice score) with TernausNet. Based on my results, image-based surgical technical skill assessment can be a good alternative to kinematic data-based skill assessment for RAMIS-type systems.

Related publications: [RNE1, RNE2, RNE3, RNE4, RNE5]

Thesis group 2

In my second Thesis group, I proved the correlation between objectively measured image and force data and surgical non-technical skills in robotic surgery.

Thesis 2/I: I proposed a methodology for autonomous non-technical skill assessment and workload evaluation in laparoscopic cholecystectomy training. The training was studied with workload assessment (SURG-TLX). Laparoscopic phantom training and workflow were introduced to simulate stressful situations during surgery (bleeding, time-critical reaction, distractions, physical demands). Statistical tests showed significant differences between the two groups (medical professionals and control group) in the case of mental demands, physical demands and situational stress ($p<0.001$, 95 % CI). Learning curve in task complexity resulted significant difference between the first and the second trials ($p<0.05$, 95 % CI).

Thesis 2/II: It has been shown that there are non-technical skills of the surgeon which can be classified based on sensory data (image and force) during the surgeon's training, based on a Fully Convolutional Neural Network (85 % accuracy). This finding can lead to objective and autonomous surgical non-technical skill assessment, and the proposed training environment can be suitable for personalized training as well.

Related publications: [RNE6, RNE7, RNE8, RNE9]

Thesis group 3

I proposed a framework for Optical Flow-based surgical skill assessment in the case of autonomous endoscope motion to assess surgical skills while ECM automation is employed.

Thesis 3/I: An accurate OF-based surgical skill assessment algorithm was presented; it outperformed the state of the art, and I proved its applicability in RAMIS skill assessment. The results were tested on the JIGSAWS dataset: in ST, NP, and KT tasks it provided 81.89 %, 84.23 % and 83.54 % accuracy, respectively, with ResNet classification.

Thesis 3/II: I proposed an autonomous endoscopic camera motion algorithm based on visual servoing, tested on the dVSS. The method could successfully compensate the distances between Patient Side Manipulators and the endoscopic camera with a marker-based approach.

Thesis 3/III: An OF ego-motion compensation method was developed, to extract only surgical tool motions in the visual scene. This framework can be employed with future safety examinations, where the motion of the endoscope is automated, while the skills of the surgeon are being assessed.

Related publications: [RNE10, RNE11, RNE12, RNE13, RNE14, RNE15]

Further publications related to the Ph.D. Thesis and the accompanying research work: [RNENR1, RNENR2, RNENR3, RNENR4, RNENR5, RNENR6, RNENR7]

5 Conclusion and practical applicability

Surgical skill assessment can improve training and patient outcome. In this dissertation, automated skill assessment in Robot-Assisted and manual MIS was studied. Endoscopic image data-based surgical technical skill assessment was shown, additional sensor- and motion analysis-based non-technical surgical skill assessment was proposed, and finally, a framework for image-based skill assessment – in the case of autonomous camera motion – was introduced. Technical robotic surgical skills were identified, and the possibility to autonomously measure these skills. For this, a 2D endoscopic image data-based skill classification method was introduced, where kinematic data was generated based on the image features. Its applicability in skill classification was examined as well. Non-technical surgical skills are also possible to estimate objectively with kinematic, endoscopic, or additional sensor-based autonomous approaches. I introduced the most important non-technical skills in RAMIS, and the possibilities to assess these skills manually and autonomously. I showed that with Artificial Intelligence methods, certain non-technical skills can be classified autonomously. Finally, a method for autonomous camera motion for the dVSS was introduced, and I proposed an OF-based method for surgical skill assessment which can be utilized when automation is employed, furthermore, I proposed an OF ego-motion filter algorithm to extract only the surgical tool motion. My work was based on human surgical data and experiments, with a market leader RAMIS system's research platform. In my thesis project, I aimed to develop widely applicable benchmarks, which can help not only the work of the RAMIS community, but surgical training as well.

My thesis work has the potential to contribute significantly to the field of autonomous skill assessment and training for surgical residents. Following my doctoral research, my focus shifted towards assessing both technical and non-technical skills among surgeon residents and console surgeons using the da Vinci Xi surgical system, incorporating physiological signal analysis into the evaluation process. This collaborative effort involved Semmelweis University (Budapest, Hungary) and Queen's University (Kingston, Canada). Surgical skill assessment plays a vital role in the emerging field of Surgical Data Science (SDS), and I have been applying my expertise to SDS studies at the Austrian Center for Medical Innovation and Technology (ACMIT). My research findings have also been integrated into various projects, such as "Robot-enhanced skill improvement and assessment in minimally invasive surgery" (OTKA PD116121, 2015–2017) and "Research on Localization and Object Detection Based on Heterogeneous Sensors" (GINOP-2.2.1-15-2017-00097, 2018–2021). Additionally, my work in the domain of DVRK-related research has been incorporated into `irob-saf`, a framework for robotic surgery automation. In my role as an assistant lecturer at the John von Neumann Faculty of Informatics, Óbuda University, I have had the privilege of sharing my research insights with BSc and MSc students. This experience allowed me to collaborate closely with talented students from institutions such as Óbuda University, Eötvös Loránd University, Budapest University of Technology and Economics and Pázmány Péter Catholic University. Looking ahead, my future projects involve non-technical skill assessment with Hungarian console surgeons in clinical settings and the development of an autonomous Surgical Process Modeling system to support clinical decision-making.

References

[1] A. Takács, D. A. Nagy, I. Rudas, and T. Haidegger, "Origins of Surgical Robotics: From Space to the Operating Room," *Acta Polytechnica Hungarica*, vol. 13, pp. 13–30, 2016.

[2] G. Fichtinger, J. Troccaz, and T. Haidegger, "Image-guided interventional robotics: Lost in translation?" *Proceedings of the IEEE*, vol. 110, no. 7, pp. 932–950, 2022.

[3] J. Chen, N. Cheng, G. Cacciamani, P. Oh, M. Lin-Brande, D. Remulla, I. S. Gill, and A. J. Hung, "Objective assessment of robotic surgical technical skill: a systematic review," *The Journal of urology*, vol. 201, no. 3, pp. 461–469, 2019.

[4] J. Heemskerk, H. R. Zandbergen, S. W. M. Keet, I. Martijnse, G. van Montfort, R. J. A. Peters, V. Svircevic, R. A. Bouwman, C. G. M. I. Baeten, and N. D. Bouvy, "Relax, it's just laparoscopy! A prospective randomized trial on heart rate variability of the surgeon in robot-assisted versus conventional laparoscopic cholecystectomy," *Digestive Surgery*, vol. 31, no. 3, pp. 225–232, 2014.

[5] D. Stefanidis, F. Wang, J. R. Korndorffer, J. B. Dunne, and D. J. Scott, "Robotic assistance improves intracorporeal suturing performance and safety in the operating room while decreasing operator workload," *Surgical Endoscopy*, vol. 24, no. 2, pp. 377–382, Feb. 2010.

[6] K. Cleary and C. Nguyen, "State of the Art in Surgical Robotics: Clinical Applications and Technology Challenges," *Computer Aided Surgery*, vol. 6, no. 6, pp. 312–328, Jan. 2001.

[7] S. G. Hart, "NASA-Task Load Index (NASA-TLX); 20 Years Later," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, pp. 904–908, Oct. 2006.

[8] B. Gibaud, G. Forestier, C. Feldmann, G. Ferrigno, P. Gonçalves, T. Haidegger, C. Julliard, D. Katić, H. Kenngott, L. Maier-Hein, K. März, E. de Momi, D. Á. Nagy, H. Nakawala, J. Neumann, T. Neumuth, J. Rojas Balderrama, S. Speidel, M. Wagner, and P. Jannin, "Toward a standard ontology of surgical process models," *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 9, pp. 1397–1408, Sep. 2018.

[9] S. Yule and S. Paterson-Brown, "Surgeons' Non-technical Skills," *Surgical Clinics*, vol. 92, no. 1, pp. 37–50, Feb. 2012.

[10] L. Hull, S. Arora, R. Aggarwal, A. Darzi, C. Vincent, and N. Sevdalis, "The Impact of Nontechnical Skills on Technical Performance in Surgery: A Systematic Review," *Journal of the American College of Surgeons*, vol. 214, no. 2, pp. 214–230, Feb. 2012.

[11] J. W. Collins, P. Dell'Oglio, A. J. Hung, and N. R. Brook, "The Importance of Technical and Non-technical Skills in Robotic Surgery Training," *European Urology Focus*, vol. 4, no. 5, pp. 674–676, Sep. 2018.

[12] C. Urban, P. Galambos, G. Györök, and T. Haidegger, "Simulated Medical Ultrasound Trainers A Review of Solutions and Applications," *Acta Polytechnica Hungarica*, vol. 15, no. 7, pp. 111–131, 2018.

[13] J. Collins and P. Wisz, "Training in robotic surgery, replicating the airline industry. How far have we come?" *World Journal of Urology*, vol. 38, no. 7, pp. 1645–1651, 2020.

[14] R. Sánchez, O. Rodríguez, J. Rosciano, L. Vegas, V. Bond, A. Rojas, and A. Sanchez-Ismayel, "Robotic surgery training: Construct validity of Global Evaluative Assessment of Robotic Skills (GEARS)," *Journal of Robotic Surgery*, vol. 10, no. 3, pp. 227–231, Sep. 2016.

[15] T. Haidegger, "Autonomy for Surgical Robots: Concepts and Paradigms," *IEEE Transactions on Medical Robotics and Bionics*, vol. 1, no. 2, pp. 65–76, May 2019.

[16] L. Zhou and M. Kaess, "An Efficient and Accurate Algorithm for the Perspecitve-n-Point Problem," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nov. 2019, pp. 6245–6252.

[17] E. Colleoni, P. Edwards, and D. Stoyanov, "Synthetic and real inputs for tool segmentation in robotic surgery," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2020, pp. 700–710.

[18] Y. Beck, T. Herman, M. Brozgol, N. Giladi, A. Mirelman, and J. M. Hausdorff, "Sparc: a new approach to quantifying gait smoothness in patients with parkinson's disease," *Journal of neuroengineering and rehabilitation*, vol. 15, no. 1, pp. 1–9, 2018.

[19] S. Singh, J. Bible, Z. Liu, Z. Zhang, and R. Singapogu, "Motion Smoothness Metrics for Cannulation Skill Assessment: What Factors Matter?" *Frontiers in Robotics and AI*, vol. 8, 2021.

[20] Y. Gao, S. S. Vedula, C. E. Reiley, N. Ahmidi, B. Varadarajan, H. C. Lin, L. Tao, L. Zappella, B. Béjar, D. D. Yuh *et al.*, "Jhu-isi gesture and skill assessment working set (jigsaws): A surgical activity dataset for human motion modeling," in *MICCAI workshop: M2cai*, vol. 3, no. 3, 2014, pp. 1–10.

[21] M. Petković, S. Džeroski, and D. Kocev, "Feature Ranking for Hierarchical Multi-Label Classification with Tree Ensemble Methods," *Acta Polytechnica Hungarica*, vol. 17, pp. 129–148, Jan. 2020.

[22] A. A. Shvets, A. Rakhlin, A. A. Kalinin, and V. I. Iglovikov, "Automatic Instrument Segmentation in Robot-Assisted Surgery using Deep Learning," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Dec. 2018, pp. 624–628.

[23] A. Chaurasia and E. Culurciello, "Linknet: Exploiting encoder representations for efficient semantic segmentation," in *2017*

IEEE Visual Communications and Image Processing (VCIP), 2017, pp. 1–4.

[24] V. Iglovikov and A. A. Shvets, “TernausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation,” *arXiv: Computer Vision and Pattern Recognition*, pp. 1–5, Jan. 2018.

[25] M. R. Wilson, J. M. Poolton, N. Malhotra, K. Ngo, E. Bright, and R. S. W. Masters, “Development and validation of a surgical workload measure: The surgery task load index (SURG-TLX),” *World Journal of Surgery*, vol. 35, no. 9, pp. 1961–1969, Sep. 2011.

[26] A. Lukezic, T. Vojir, L. Cehovin Zajc, J. Matas, and M. Kristan, “Discriminative correlation filter with channel and spatial reliability,” in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2017, pp. 6309–6318.

[27] Z. Wang, W. Yan, and T. Oates, “Time series classification from scratch with deep neural networks: A strong baseline,” in *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 1578–1585.

[28] S. Hutchinson, G. Hager, and P. Corke, “A tutorial on visual servo control,” *IEEE Transactions on Robotics and Automation*, vol. 12, no. 5, pp. 651–670, Oct. 1996.

[29] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights into Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018.

[30] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[31] R. P. Monti, S. Tootoonian, and R. Cao, “Avoiding degradation in deep feed-forward networks by phasing out skip-connections,” in *Artificial Neural Networks and Machine Learning–ICANN*

2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27. Springer, 2018, pp. 447–456.

[32] Q. V. Le *et al.*, "A tutorial on deep learning part 2: Autoencoders, convolutional neural networks and recurrent neural networks," *Google Brain*, vol. 20, pp. 1–20, 2015.

[33] A. Zia, Y. Sharma, V. Bettadapura, E. L. Sarin, T. Ploetz, M. A. Clements, and I. Essa, "Automated video-based assessment of surgical skills for training and evaluation in medical schools," *International Journal of Computer Assisted Radiology and Surgery*, vol. 11, no. 9, pp. 1623–1636, Sep. 2016.

Publications related to the Thesis

[RNE1] R. Nagyné Elek and T. Haidegger, "Robot-assisted minimally invasive surgical skill assessment—manual and automated platforms," *Acta Polytechnica Hungarica*, vol. 16, no. 8, pp. 141–169, 2019.

[RNE2] D. El-Saig, R. Nagyné Elek, and T. Haidegger, "A Graphical Tool for Parsing and Inspecting Surgical Robotic Datasets," in *2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI)*, 2018, pp. 131–136.

[RNE3] R. Nagyné Elek and T. Haidegger, "Endoszkópos kameraképek feldolgozásának lehetőségei a robotsebészetben," *IME: Informatika és Menedzsment az Egészségügyben*, vol. 17, no. 4, pp. 49–54, 2018.

[RNE4] D. Papp, R. Nagyné Elek, and T. Haidegger, "Surgical tool segmentation on the jigsaws dataset for autonomous image-based skill assessment," *IEEE 10th Jubilee International Conference on Computational Cybernetics and Cyber-Medical Systems (ICCC)*, pp. 49–56, 2022.

[RNE5] R. Nagyné Elek and T. Haidegger, "Towards autonomous endoscopic image-based surgical skill assessment: Articulated tool pose estimation," *IEEE 10th Jubilee International Conference on Computational Cybernetics and Cyber-Medical Systems (ICCC)*, pp. 35–42, 2022.

[RNE6] R. Nagyné Elek and T. Haidegger, "Non-technical skill assessment and mental load evaluation in robot-assisted minimally invasive surgery," *Sensors*, vol. 21, no. 8, p. 2666, 2021.

[RNE7] R. Nagyné Elek, B. Mach, K. Móga, A. Ládi, and T. Haidegger, "Autonomous non-technical surgical skill assessment and workload analysis in laparoscopic cholecystectomy training," *2021 IEEE 25th International Conference on Intelligent Engineering Systems (INES)*, pp. 33–39, 2022.

[RNE8] R. Nagyné Elek and T. Haidegger, "Future Trend: AI x Robot," *Artificial Intelligence in Surgery – Recent Advances and Future (Springer Nature), editor: Ken Masamune*, p. 22 (in press), 2023.

[RNE9] R. Nagyné Elek and T. Haidegger, "Next in surgical data science: Autonomous non-technical skill assessment in minimally invasive surgery training," *Journal of Clinical Medicine*, vol. 11, no. 24, p. 7533, 2022.

[RNE10] R. Elek, T. Dániel Nagy, D. Á. Nagy, G. Kronreif, I. J. Rudas, and T. Haidegger, "Recent trends in automating robotic surgery," in *2016 IEEE 20th Jubilee International Conference on Intelligent Engineering Systems (INES)*, 2016, pp. 27–32.

[RNE11] C. Molnár, T. D. Nagy, R. Nagyné Elek, and T. Haidegger, "Visual servoing-based camera control for the da Vinci Surgical System," in *2020 IEEE 18th International Symposium on Intelligent Systems and Informatics (SISY)*, 2020, pp. 107–112.

[RNE12] C. Molnár, R. Nagyné Elek, T. D. Nagy, R. Kiss, and T. Haidegger, "Vizuális kameravezérlés da Vinci sebészeti robotrendszerre: Visual servoing for da Vinci surgical system," *Nemzetközi Gépészeti Konferencia – OGÉT*, pp. 329–332, 2020.

[RNE13] G. Lajkó, R. Nagyné Elek, and T. Haidegger, "Surgical Skill Assessment Automation Based on Sparse Optical Flow Data," in *2021 IEEE 25th International Conference on Intelligent Engineering Systems (INES)*, 2021, pp. 201–208.

[RNE14] G. Lajkó, R. Nagyné Elek, and T. Haidegger, “Endoscopic Image-Based Skill Assessment in Robot-Assisted Minimally Invasive Surgery,” *Sensors*, vol. 21, no. 16, p. 5412, 2021.

[RNE15] R. Elek, A. Károly, T. Haidegger, and P. Galambos, “Towards Optical Flow Ego-motion Compensation for Moving Object Segmentation,” in *International Conference on Robotics, Computer Vision and Intelligent Systems*, 2022, pp. 114–120.

Other publications

[RNENR1] R. Elek, T. D. Nagy, D. Á. Nagy, T. Garamvölgyi, B. Takács, P. Galambos, J. K. Tar, I. J. Rudas, and T. Haidegger, “Towards surgical subtask automation — Blunt dissection,” in *2017 IEEE 21st International Conference on Intelligent Engineering Systems (INES)*, 2017, pp. 253–258.

[RNENR2] R. Elek, T. D. Nagy, D. Á. Nagy, B. Takács, P. Galambos, I. Rudas, and T. Haidegger, “Robotic platforms for ultrasound diagnostics and treatment,” in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 1752–1757.

[RNENR3] A. I. Károly, R. N. Elek, T. Haidegger, K. Széll, and P. Galambos, “Optical flow-based segmentation of moving objects for mobile robot navigation using pre-trained deep learning models*,” in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019, pp. 3080–3086.

[RNENR4] A. I. Károly, R. Nagyné Elek, T. Haidegger, and P. Galambos, “Moving Obstacle Segmentation with an

Optical Flow-based DNN: An Implementation Case Study," in *2021 IEEE 25th International Conference on Intelligent Engineering Systems (INES)*, 2021, pp. 189–194.

[RNENR5] D. Á. Nagy, T. D. Nagy, R. Elek, I. Rudas, and T. Haidegger, "Ontology Based Surgical Subtask Automation," in *International Conference on Robotics and Automation 2017*, Singapore, 2017.

[RNENR6] D. Á. Nagy, T. D. Nagy, R. Elek, I. J. Rudas, and T. Haidegger, "Ontology-Based Surgical Subtask Automation, Automating Blunt Dissection," *Journal of Medical Robotics Research*, vol. 3, no. 3, p. 1841005, 2018.

[RNENR7] K. Takács, R. N. Elek, and T. Haidegger, "Image processing-based methods to improve the robustness of robotic gripping," in *2022 IEEE 22nd International Symposium on Computational Intelligence and Informatics and 8th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Science and Robotics (CINTI-MACRo)*. IEEE, 2022, pp. 345–350.