

Adaptive Neuro-fuzzy Inference System for Automated Skill Assessment in Robot-Assisted Minimally Invasive Surgery

Kristóf Takács

*Antal Bejczy Center of Intelligent Robotics, EKIK & Doctoral School of Applied Informatics and Applied Mathematics
Óbuda University
Budapest, Hungary
kristof.takacs@irob.uni-obuda.hu*

Tamás Haidegger

*Antal Bejczy Center of Intelligent Robotics, EKIK & John von Neumann Faculty of Informatics
Óbuda University
Budapest, Hungary
tamas.haidegger@irob.uni-obuda.hu*

Abstract—The new kinds of minimally invasive surgical tools and methods require new and complex skillsets from the surgeons, thus the objective skill-based training became essential. This new domain can greatly rely on the automated assessment of surgical skills. There are many devices and approaches to measure different aspects of surgical task accution, however, the results are often hard to interpret. This paper describes an adaptive neuro-fuzzy inference system for the classification of the performance of subjects using a skill assessment device designed for psychomotor skill training for robotic surgical procedures (the FRS Dome). The FRS Dome offers 7 independent surgical tasks, this paper focuses on two of them, since the rest were often impossible to perform for novices. The fuzzy-systems for the two tasks were designed based on 27 performances of subjects with varying skill-levels, the system is capable of scoring the tasks separately, and also scoring the whole performance on all the tasks together on 1–3 scales. In general, the neuro-fuzzy system is capable of optimizing the classification based on future measurements, thus our method will be able to tune the classification of the rest of the tasks based on upcoming trial results.

Index Terms—Neuro-fuzzy inference system, surgical skill assessment, RAMIS

I. INTRODUCTION

The most common spread of laparoscopic surgery (also referred to as Minimal Invasive Surgery, MIS) arguably opened a new chapter in modern medicine. During these procedures, the surgeon uses special, narrow shafted surgical instruments and an endoscopic camera inserted into the body through small incisions. In addition, robots and different precision

Kristóf Takács acknowledges the financial support of Óbuda University – Doctoral School of Applied Informatics and Applied Mathematics.

This work was supported by the UNKP-20-4 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

The publication of this article has been supported by the Robotics Special College via the "NTP-SZKOLL-20-0043 Talent management and community building in the Robotics Special College of Óbuda University" project.

The publication of this article has been supported by the Robotics Special College via the "NTP-SZKOLL-20-0043 Talent management and community building in the Robotics Special College of Óbuda University" project.

T. Haidegger is a Bolyai Fellow of the Hungarian Academy of Sciences, his research is partially supported by the 2019-1.3.1-KK-2019-00007 project.

mechatronics in the operating room are also providing additional support, opening to Robot Assisted Minimal Invasive Surgery (RAMIS) [1]. Generally, these new approaches result in less tissue-trauma and shorter recovery time, however, the surgeons need to learn the controls of the robots and different machines too, beside handling classical and laparoscopic surgical instruments. Furthermore, during MIS and RAMIS procedures, the feedbacks that the surgeon receives changed dramatically; instead of seeing directly the operating area and the instruments, the endoscopic camera provides the only visual feedback, with serious inherent limitations, and there is typically no haptic feedback during RAMIS procedures.

As a result of the fast improvement of surgical methods and devices, surgical skill training and skill assessment became even more important—but also even more difficult than before [2], [3]. Along with the recent development of surgery itself, new technologies offer new options for surgical skill training too, however the breakthrough in the field of surgical skill assessment is yet to be seen. Simulators and box-trainers are good examples of new and constantly developing training methods. Low- and high-fidelity phantom or virtual models can be created to train different surgical tasks, hand-eye coordination, basic movements or complete surgical procedures [4], [5]. The most widely accepted complete training curricula nowadays for MIS and RAMIS are the Fundamentals of Laparoscopic Surgery (FLS) and the Fundamentals of Robotic Surgery (FRS) programs, respectively [6], [7], [8].

Surgical skill training platforms are essential for surgeons to safely prepare for real operations, however, the traditional peer-assessment system can not ensure objective measurement of the acquired skills. The first step towards objectivity was the introduction of manual skill assessing tables. These spreadsheets (e.g., GOALS [9] and GEARS [10]) request the reviewer to evaluate the trainees by giving scores on the pre-defined aspects of the performance, thus documenting the process, and making it less subjective. However, both the box-trainers and virtual simulators have the potential to provide automated skill assessment, meaning that the evaluation may

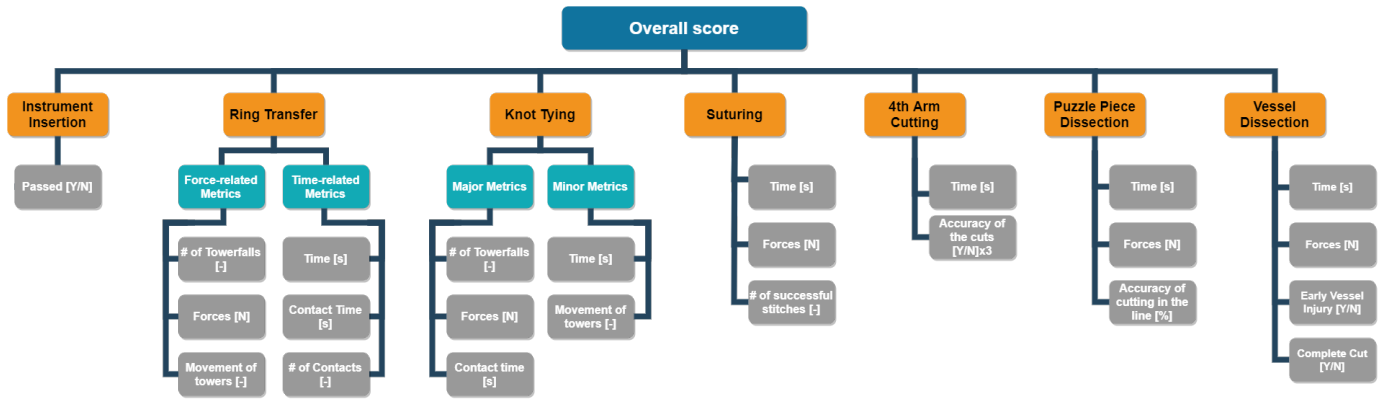


Fig. 1. The measured and calculated metrics for the tasks of the FRS Dome, 22 different metrics or errors are monitored during the 6 psychomotor skill training tasks (the 1st task is introductory, not a surgical task).



Fig. 2. The original FRS Dome [11].

be carried out by an algorithm, based on sensor data, image processing and other metrics.

In an earlier paper, we presented a psychomotor skill training device – the sensorized FRS Dome – capable of automated skill assessment [12]. The original FRS Dome (Fig. 2) was designed by a group of professionals for the FRS curriculum with the goal of creating a standalone, platform-independent phantom for psychomotor skill training [7]. Surgeons can practice some basic and common tasks of RAMIS through 7 exercises on the device, however the original phantom was not capable of automated scoring. Takács et al. in [12] described the sensorization process, and the initial results, proving the correlation between the measured metrics and the skill-level of the subjects. However, the measured outcome were sometimes only raw sensor-data or computed values that were relevant results (i.e., correlated with the skill-level), but were hard to interpret from the medical point of view.

This paper describes a neuro-fuzzy inference system that calculates scores for the 7 tasks (and an overall score too), based on the metrics measured and calculated by the integrated sensors (3 force sensors, contact sensors, camera) and the microcontroller [12]. After the description of the basic fuzzy-

system, an adaptive neuro-fuzzy inference system (ANFIS) is presented, that will help optimizing the boundaries of the classifications in long term, based on future measurements of surgeons with known skill-levels.

II. METHODS

A. Metrics of the FRS tasks

The two main area of application of fuzzy systems are control theory and classification problems, this paper covers the latter [13] [14]. The purpose of the implementation of a fuzzy system was to realize a classification of the human performances on the FRS Dome, i.e., to provide easily interpretable scores (e.g. *bad-average-good* classification) based on the measured metrics.

The fuzzy system based classification consists of two main levels of approximation. The first is the level of the surgical tasks (some of them further divided into two levels), the second level is the overall scoring based on the results of all the seven individual tasks (Fig. 1). This two-level grading system has two main advantages. First, the subjects do not need to complete all the tasks to receive scores, which is important, since novice subjects often can not accomplish the more complicated tasks at all (or only with several "fatal errors"). Second, calculating an overall score in one fuzzy inference system using all the 22 measured metrics would have resulted in a too complex system, being difficult to handle, even with the modern fuzzy-toolboxes. This second argument was the reason for the further division of the tasks with 5 or more metrics too.

The measured and calculated metrics and the whole sensorization of the phantom was designed based on the official descriptions of the tasks (with the important metrics, minor and fatal errors listed for each) published and also made available on the FRS webpage [7] [15]. Fig. 1 shows the 7 FRS-tasks (orange boxes) and the measured metrics for each tasks used as inputs in the fuzzy systems (grey boxes). The first task (*Instrument Insertion*) is excluded from this research, since it does not measure technical surgical skills. The scores for the second and third tasks (*Ring Transfer* and *Knot Tying*)

are calculated using a 2-level fuzzy-tree, since 5-6 inputs are difficult to handle in a Sugeno-type fuzzy systems. The results for the rest of the tasks are obtained through general Sugeno-type fuzzy inference systems.

B. Basic principles of the fuzzy systems

In this paper, the fuzzy systems of the second and the fourth tasks (*Ring Transfer* and *Suturing*) will be explained as examples, the methods for the rest of the tasks are similar to these two tasks. The fuzzy rules, membership functions, weights and the classification of metrics were determined based on the official publications and task descriptions by the FRS group, beside consulting with a fuzzy systems expert and a laparoscopic surgeon resident.

Throughout the whole project, the MATLAB software was used (*Matlab 2020b* by *The MathWorks Inc.*), with the *Fuzzy Logic Designer* and the *Neuro-Fuzzy Designer* toolboxes. The initial data was collected from novices, medical students and laparoscopic surgeons, the methods of data collection and initial analysis are described in Takács et al. [12]. All the developed fuzzy systems use *products* as "And method", *probabilistic OR* as "Or method" and "weighted average" for defuzzification (denoted as *prod*, *probor* and *wtaver* respectively in Matlab environment).

The goal from the beginning was the development of an ANFIS, therefore the practical decision was the development of Sugeno-type fuzzy inference systems. The Matlab toolboxes support Sugeno fuzzy systems and fuzzy trees too, thus all the components and the whole fuzzy tree was built in that environment, from bottom to top.

C. Creating the fuzzy systems

Fig. 1 can be interpreted as the visualisation of the structure of the fuzzy tree too. The scoring of the *Ring Transfer* task is carried out by a simple fuzzy tree, the *Force-related Metrics* and the *Time-related Metrics* give a sub-score each, these are the inputs for the fuzzy system calculating the final score for the task. The membership functions are triangular, the scores of the two sub-systems are counting with the same weight in the final score, thus the surface of the fuzzy tree function (the output value as a function of the two inputs) is symmetric for the two inputs.

The scoring system of the *Knot tying* task is similar to the one described above (with different metrics as inputs). The only difference being that the surface shown on Fig. 3 would be different, since the first group of metrics is more important than the second, thus the surface is asymmetric.

The rest of the task-related fuzzy systems are simpler, since they are not fuzzy-trees, only simple Sugeno type fuzzy inference systems. The classifying fuzzy systems are similar to one level of the previously described systems, all having one output (the final score of the task on the *bad-average-good* scale), and as many inputs as shown on Fig. 1.

The "upper-level" fuzzy system has 6 inputs and 1 output, giving a final score for the subject's performance based on all the tasks performed on the phantom. Since the initial goal

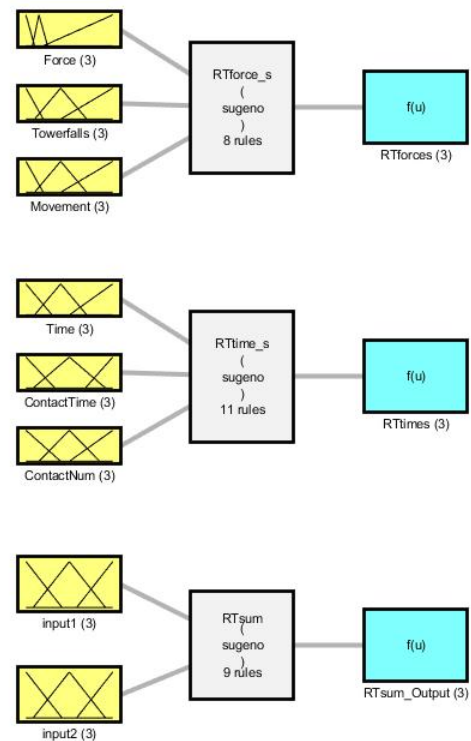


Fig. 3. The Sugeno fuzzy systems and the fuzzy-tree for calculating the final score for the *Ring Transfer* task. The two inputs for the fuzzy-tree are the outputs of the two upper fuzzy systems. The graph shows the generated function of the fuzzy-tree based on the given rules.

of the project was to develop a device capable of automated skill assessment for robotic surgery trainees, the final score is a grade on the scale of 1–3. The official descriptions from the FRS group does not distinguish more and less important tasks, thus this fuzzy system – despite the high number of inputs – is not complicated. All the inputs share the same value set (i.e., the outputs of the lower-level fuzzy systems are all scores on a 1–3 scale), and all the inputs should count with the same weights, thus the rules of the Sugeno type fuzzy system are easy to generate.

D. The adaptive neuro-fuzzy inference system

The neuro-fuzzy inference systems are typically used to replace huge rulesets with significantly less rules and an adaptive learning algorithm. In the case of ANFIS, Matlab requires the same number of rules and output membership functions, thus only 3 fuzzy rules can be kept at each task.

Fig. 4 shows the second group of metrics for the *Knot Tying* task (denoted as "Minor Metrics" on Fig. 1), since this 2 input-1 output fuzzy system is the easiest to visualize. The training dataset was the set of measurements of subjects with known skill-levels and a simulated dataset was used for testing. The training dataset, however, was not big enough (less entries than tunable parameters), which might explain that the trained system gives lower results in most cases.

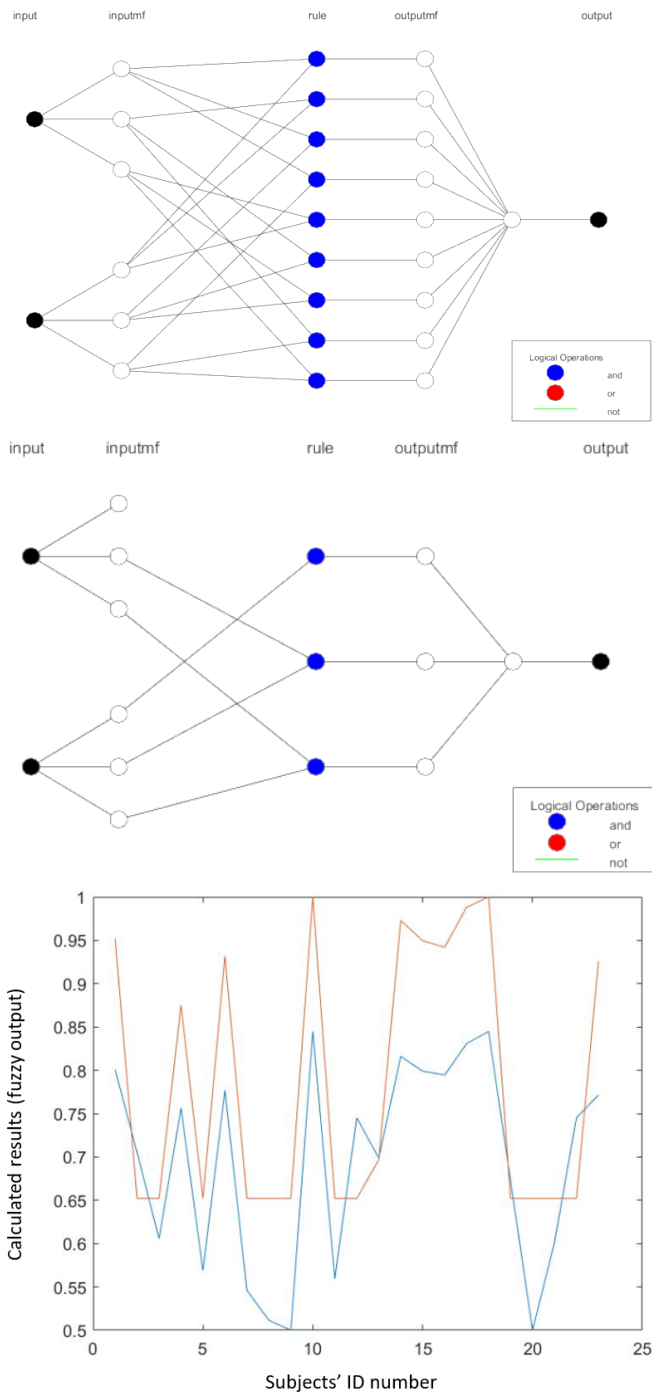


Fig. 4. Visualization of the original fuzzy system and the initial fuzzy system used in ANFIS for the "Minor metrics" of the Knot Typing task (top and middle respectively), and the results for a set of real performances calculated by the original (orange) and the trained (blue) systems.

III. RESULTS AND DISCUSSION

The task-level fuzzy systems were constructed to handle lists of results, making it easy to calculate the scores of several performances at once (since the sensorized FRS Dome described in [12] produces spreadsheets separately for each task, each line representing one subject's performance in each

spreadsheet). With the options of Matlab, C-code can be generated from all the software created in Matlab, which can be integrated in most kind of projects, e.g., the Labview interface that handles the sensorized phantom.

Due to the availability of professional robotic surgeons in Hungary, the general 3 groups of "novices", "intermediates" and "experts" could not be set up and compared with statistical methods, thus the learning curves of novices and intermediates were examined. Previously in [12], the same method was used to show the correlation between the measured metrics and the level of expertise, thus the purpose of this analysis was only the inspection of the fuzzy-system. It was found that the 3rd performances of the subjects resulted in the same, or better score in all 6 cases regarding the first two tasks, however the rest of the tasks could not be examined due to the lack of data, since novices could not perform them without fatal errors.

IV. CONCLUSION

This paper presented a set of Sugeno type fuzzy systems (a fuzzy tree) that are capable of assigning scores in an objective and automated way to robotic surgery psychomotor skill training tasks. The task-level fuzzy systems gave scores on each of the 7 FRS-tasks separately, while the top-level fuzzy system provided a grade on a (*bad-average-good*) scale based on all the measured metrics during the performance of the 7 tasks. In addition to the original fuzzy systems, adaptive neuro-fuzzy inference systems were developed for each task, that may give optimized results based on an appropriate number of measured performances of subjects with known robotic surgery skill level.

REFERENCES

- [1] T. Haidegger, "Autonomy for surgical robots: Concepts and paradigms," *IEEE Transactions on Medical Robotics and Bionics*, vol. 1, no. 2, pp. 65–76, 2019.
- [2] 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.
- [3] —, "Non-technical skill assessment and mental load evaluation in robot-assisted minimally invasive surgery," *Sensors*, vol. 21, no. 8, p. 2666, 2021.
- [4] B. Schwab, E. Hungness, K. A. Barsness, and W. C. McGaghie, "The role of simulation in surgical education," *Journal of Laparoendoscopic & Advanced Surgical Techniques*, vol. 27, no. 5, pp. 450–454, 2017.
- [5] I. Nigicsér, B. Szabó, L. Jaksa, D. Á. Nagy, T. Garamvölgyi, S. Barcza, P. Galambos, and T. Haidegger, *CogInfoCom-Driven Surgical Skill Training and Assessment*. Springer, Cham, 2018.
- [6] J. H. Peters, G. M. Fried, L. L. Swannstrom, N. J. Soper, L. F. Sillin, B. Schirmer, K. Hoffman, S. F. Committee *et al.*, "Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery," *Surgery*, vol. 135, no. 1, pp. 21–27, 2004.
- [7] R. Smith, V. Patel, and R. Satava, "Fundamentals of robotic surgery: a course of basic robotic surgery skills based upon a 14-society consensus template of outcomes measures and curriculum development," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 10, no. 3, pp. 379–384, 2014.
- [8] D. Herron and M. Marohn, "A consensus document on robotic surgery," *Surgical endoscopy*, vol. 22, no. 2, pp. 313–325, 2008.
- [9] M. C. Vassiliou, L. S. Feldman, C. G. Andrew, S. Bergman, K. Lefondré, D. Stanbridge, and G. M. Fried, "A global assessment tool for evaluation of intraoperative laparoscopic skills," *The American journal of surgery*, vol. 190, no. 1, pp. 107–113, 2005.

- [10] A. C. Goh, D. W. Goldfarb, J. C. Sander, B. J. Miles, and B. J. Dunkin, "Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills," *The Journal of urology*, vol. 187, no. 1, pp. 247–252, 2012.
- [11] B. T. Carpenter and C. P. Sundaram, "Training the next generation of surgeons in robotic surgery," *Robotic surgery (Auckland)*, vol. 4, pp. 39–44, 2017.
- [12] K. Takács, K. Móga, and T. Haidegger, "Sensorized psychomotor skill assessment platform built on a robotic surgery phantom," in *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2020, pp. 95–100.
- [13] A. d. Amo, J. Montero, G. Biging, and V. Cutello, "Fuzzy classification systems," *European Journal of Operational Research*, vol. 156, no. 2, pp. 495–507, 2004.
- [14] R.-E. Precup, L. Kovács, T. Haidegger, S. Preitl, A. Kovács, B. Benyó, E. Borbély, and Z. Benyó, "Time delay compensation by fuzzy control in the case of master-slave telesurgery," in *2011 6th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI)*. IEEE, 2011, pp. 305–310.
- [15] F. of Robotic Surgery, "Fundamentals of robotic surgery," Available: <http://www.frsurgery.org>. [Accessed: 2021.05.30.].

