Towards the Development of an Automated Assessment System for the Fundamentals of Laparoscopic Surgery Tests

Mohsen Mohaidat[1], Fatemeh Rashidi Fathabadi[1], Koloud N. Alkhamaiseh[2], Janos Grantner[1], Saad A Shebrain[3], Ikhlas Abdel-Qader[1]

[1] Electrical & Computer Engineering Department; Western Michigan University, Kalamazoo, 49008, Michigan, USA, {mohsen.mohaidat, janos.grantner, fatemeh.rashidifathabadi, ikhlas.abdelqader}@wmich.edu

[2] Department of Computer Science; Michigan Technological University, Houghton, 49931, Michigan, USA, kalkhama@mtu.edu

[3] Department of Surgery; Western Michigan University, Homer Stryker MD School of Medicine, Kalamazoo, 49008, Michigan, USA, saad.shebrain@wmed.edu

Abstract: The objective of this paper is to provide an overview of projects carried out in the framework of a research collaboration between the Department of Electrical and Computer Engineering and the Department of Surgery, in automated performance assessment for laparoscopic surgery training and testing. This paper focuses on describing the development of deep learning algorithms for object detection and tracking along with computer vision algorithms for performance assessment of Fundamentals of Laparoscopic Surgery (FLS) tests. The proposed automated assessment systems are based on quantitative measurements and expert knowledge using fuzzy logic. The Intelligent Box-Trainer System (IBTS) was used to create videos of several FLS tasks with the assistance of the medical school's surgery residents. Deep Learning (DL) models were developed and trained for three main tests of FLS: Precision Cutting, Peg Transfer, and Suturing. We placed our deep learning models in a publicly accessible database over the internet. The precision of our results compares favorably with other published work and with more data extracted from new videos, the fuzzy logic-based assessment system can be fine-tuned for even better performance.

Keywords: Object detection; laparoscopic surgery tools tracking; FLS tests; autonomous surgery skill assessment; deep learning models; fuzzy logic

1 Introduction

The first laparoscopic cholecystectomy was performed in 1985 [1]. Since then, many open surgeries have been converted to minimally invasive surgery (MIS). Although many limitations restrict conversions, such as peritoneal access, thermocoagulation, anesthesia, and pneumoperitoneum, MIS is desirable because its advantages include avoiding large wounds, less postoperative pain, and earlier hospital discharge. The small portholes used during laparoscopic surgery (LS) lead to improved cosmesis [2], [3]. Nevertheless, in order to perform laparoscopic surgery, surgeons must advance their hand skills as well as their skills in needle insertion, transferring, suturing and knotting procedures, and incision prior to performing actual surgery [4]. The Society of American Gastrointestinal and Endoscopic Surgeons initiated the Fundamentals of Laparoscopic Surgery (FLS) program in the late 1990s, a comprehensive program aimed at teaching mental and psychomotor characteristics needed for laparoscopic surgery, as well as established an assessment mechanism [5]. Specifically, it was designed to provide tools for teaching and assessing cognitive knowledge, technical skills, and clinical judgment related to basic laparoscopic surgery. Originally, seven FLS exercises or tests had been considered but only 5 of them have been chosen because two other ones have been proven not to provide meaningful assessment contribution. Consequently, the FLS consists of five tasks as follows: Peg Transfer, Precision Cutting, Ligating Loop, Suture with Extracorporeal Knot, and Suture with Intracorporeal Knot [5].

Surgeons are increasingly using surgical simulators as part of resident training to develop more complex surgical skills and improve surgical competency. A simulator may be classified as a physical task trainer (PTT), a virtual reality (VR) simulator [6], or an augmented reality (AR) simulator, depending on its state of development. A combination of virtual reality and synthetic tissue models is used in augmented reality. In the laparoscopic box trainer device that mimics the body cavity, traditional laparoscopic instruments are used [7]. Although the VR method potentially provides a very sophisticated experience, it is not without limitations. One of those limitations is the desired haptic feedback to track the impact of the surgeon's hand movements, which requires a variety of tunings in natural surgical environments [8]. There is also an example of human limits; the user may perceive the procedure to be similar to a video game rather than a surgical simulator [9]. Scientists have concluded that the use of video games as surgical simulators has created a hindrance [10]. Additionally, systems may produce dishonest acuity due to inaccurate habits acquired through an addiction to a virtual environment [9]. Hence, MIS training must provide a consistent, robust, and also affordable assessment system to avoid these constraints.

Our proposed methods and assessment systems have been presented and published in several articles in the field of developing autonomous laparoscopic assessment systems. We aim in this study to summarize these works since each of them addressed a particular FLS test by using a different computer vision algorithm.

All of these algorithms propose an FLS automated assessment system for three main laparoscopic exercises, i.e., Peg Transfer, Precision Cutting, and Suturing. In addition, we present those algorithms as they applied to several assessment parameters (such as Hand Motion assessment, Surgical Tools movement assessment, Upward Force, Execution Time, and Idle Time) along with the pertinent statistical results. In this paper, we evaluate the different Deep Learning models that we developed and customized and their main applications in developing an autonomous laparoscopic assessment system. This paper makes the following major contributions:

- It provides an overview of deep learning publications in laparoscopic box training. To enable surgical skill assessment, the publications are categorized according to surgical task: peg transfer, precision cutting, and intracorporeal suturing. Researchers seeking studies related to their work will find this design useful as a reference.
- It describes the publicly available surgical datasets that can be used by researchers to validate their deep learning models. These public datasets are also provided as links for readers to download them for future research.
- A technical analysis of the publications included in this survey is presented for each surgical task. A classification of the studies is presented in these comparisons, along with relevant information about the DL models used, the type of input data, the surgical procedure analyzed, and the dataset used to validate the models. Performance metrics are also presented to facilitate comparison between the different approaches presented.
- Further, this work demonstrates the benefits of introducing fuzzy logic models along with deep learning modules for FLS to analyze surgical performance assessments.

Throughout the rest of the paper, the following structure is followed. A brief overview of deep learning architecture and algorithms is provided in Section 2, along with a description of some of the most common models used in minimally invasive surgery applications. In Section 3, our proposed methodology for each currently supported FLS task is described. Section 4 describes the publicly available datasets of MIS along with our datasets and provides links for downloading those datasets. As a result of applying DL modules and fuzzy logic methods, Section 5 presents and discusses the results of each surgical exercise. The conclusion is outlined in Section 6.

2 Overview of Deep Learning Methods and the Use of Fuzzy Logic for a Decision Support System

To learn a high-level representation of the data useful to solve a particular problem, the deep learning (DL) network receives a set of inputs that are successively transformed by processing units called hidden layers. A particular weight and bias are assigned to units in each layer for them to be connected with units in adjacent layers. According to an activation function, the weighted sum of the inputs is transformed for each layer. In the next layer, the output from this function is used as an input to the subsequent unit [11]. DL techniques aim to minimize the loss function that measures the distance between the prediction of the network and the objectives from the training data by learning or adjusting the network parameters (connection weights and bias). The general architecture is illustrated in Fig. 1. It is primarily the automation of the feature extractors without the requirement of any manual design that distinguishes this generation of machine learning techniques from earlier ones. Thanks to the introduction of the backpropagation learning algorithm in the mid-1980s, DL has made tremendous progress during this period, as it was possible to calculate the contribution each parameter of the network made to the final loss from the outer layers down to the bottom layers. In the last ten years, there has been a proliferation of relatively inexpensive and powerful DL processing units, as well as an explosion in big data [12].

Many open-source frameworks and libraries are available that integrate training algorithms, complex mathematical functions, and statistical modeling required for the development of DL applications to support the implementation of these architectures. Many of these tools can be found on GitHub in the form of repositories. In our studies [7], [13]-[18], the DL was implemented to process input images obtained from several laparoscopic video recordings. The surgical tools were detected and tracked using several DL algorithms, such as YOLO, TensorFlow object detection, color detection, and centroid tracking. As part of Section 3 and Section 4, we will describe our proposed DL algorithms and data collection methods, respectively. Although DL may provide proper evaluation for some surgical skill assessments, it is not suitable for assessing certain parameters of advanced surgical skills, e.g., the upward force applied by the needle driver, or the force applied by the jaws of the grasper. So, DL in our work is primarily used for processing surgical video records and images. Fuzzy logic can be used to overcome these limitations along with measured data. Fuzzy logic can address noise problems and uncertainties, and incorporate expert knowledge in the decision process [19].

Fuzzy logic has been extensively used in decision support systems. In general, fuzzy inference systems (FIS) can be used to map an input space to an output space. A wide range of fields have benefited from the use of FIS, including control, decision support, image processing, and expert systems.

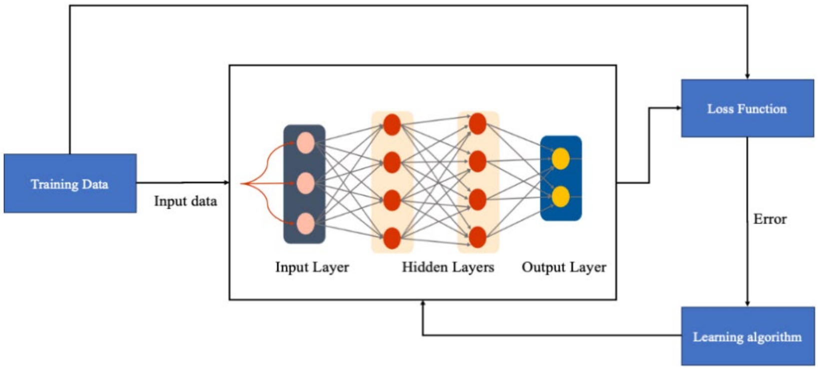


Figure 1
Training process architecture in deep learning algorithms

There are important characteristics that contribute to the strength of FIS. They can handle linguistic concepts along with quantitative measurements in the same mathematical framework. In addition, they are universal approximators capable of performing nonlinear mappings between inputs and outputs. Fuzzy linguistic terms are used to describe the fuzzy rules of the system [20]. The fuzzy rules follow the format IF (conditions) THEN (actions). The FIS is made up of the main components as follows: fuzzifier, knowledge base, inference engine, and defuzzifier, as shown in Figure 2. By processing recorded videos on the execution of FLS tests along with other, optional measurement data, and building a knowledge base using expert surgeon's opinion, the supervisor surgeon can be substituted by an objective, knowledge-based decision support system.

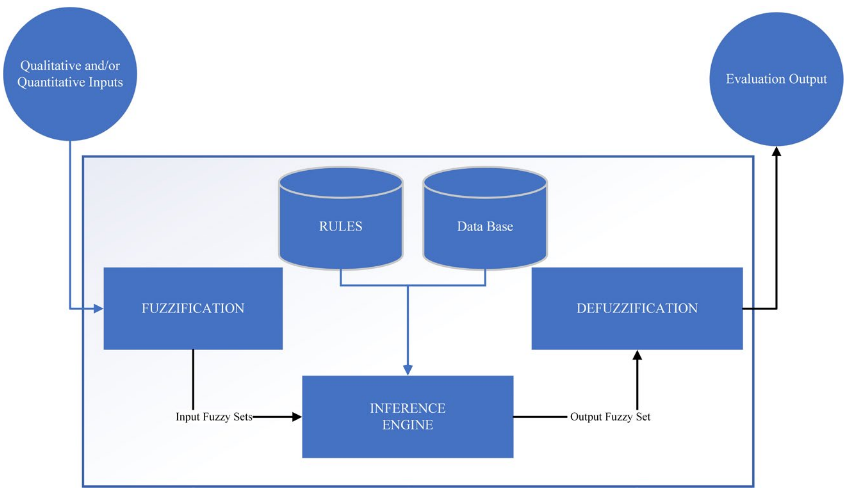


Figure 2
Fuzzy Expert System [20]

3 Applied Methodologies

Multiple experiments have been carried out in the Intelligent/Fuzzy Controllers Laboratory of WMU to validate the accuracy of our methods, both on recorded videos and with the IBTS.

This section illustrates the methodologies that were selected for each laparoscopic training exercise. In each surgical exercise, we demonstrate the DL method that has been used, the advanced surgical skills parameters that must be calculated, and the fuzzy logic methods that have been applied. Finally, we analyze and discuss the methods‘ performance for implementing an autonomous laparoscopic surgical assessment system.

3.1 Precision Cutting

One of the platforms for precision cutting exercises is composed of artificial tissue with two circles sharing the same center point. The inner circle has a radius of 2.5 centimeters, and the outer circle has a radius of 3.0 centimeters, respectively. Cuts by the scissors should be kept within the space between the two circles. After the test has started, the movement of any object outside the outer circle circumference or inside the inner circle circumference, or the cutting of a circle line, is considered an error. For this test, the laparoscopic surgical skills assessment is based on surgical tool detection and motion analysis. Fig. 3 depicts the platform along with the scissors.

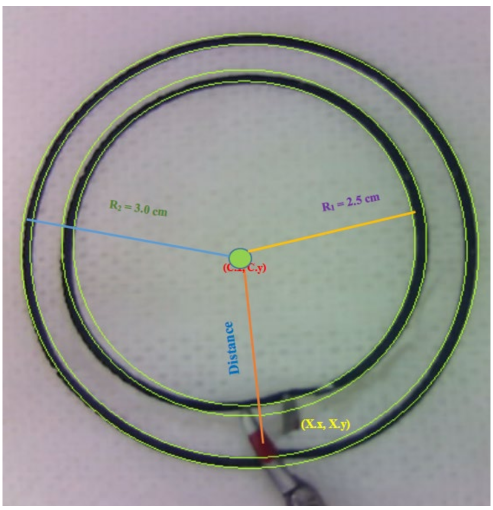


Figure 3
Illustration of artificial tissue used for pattern-cutting exercises

3.1.1 Methods

In this FLS task, five main image processing methods were used (i.e., Blob Detection [21], tip tracking [17], error calculation, fuzzy assessment, and canny edge detector [22]) which were applied to the detection, tracking, error estimation, and assessment tasks. To identify the surgical scissors, the scissor's tip was painted a red color to use a color detection method. It was established that errors may occur as a result of the change in position of the scissor's tip with respect to the center of the circle. Therefore, if the distance is less than 2.5 cm or greater than 3.0 cm, then an error flag will be set. Ideally, the trainee surgeon is expected to cut the tissue in the middle between the two circle lines. A fuzzy logic assessment system was proposed to address the uncertainty associated with trainee surgeon's performance. As shown in Table 3, five fuzzy IF-THEN rules that represent the input based on the error calculation method and are related to the scissors' tip distance from the center. According to the recommendations by an expert surgeon in the Homer Stryker MD School of Medicine (WMed), only five linguistic input values and three linguistic output values (LGs) are considered in the assessment system. The respective membership functions (MFs) are not given here.

Table 3
Memberships & Linguistic variables of Fuzzy evaluation for pattern cutting task [17]

LG & MF / Input	**Excellent** (70-100)	**Good** (60-80)	**Bad** (0-70)
Very Near Inner (-0.25 - 0.2 cm)			✓
Near Inner (-0.2cm _ -0.19cm)		✓	
Desired (-0.175cm _ 0.175cm)	✓		
Near Outer (0.19cm _ 0.21cm)		✓	
Very Near Outer (0.2 cm – 0.25 cm)			✓

3.2 Peg Transfer

As an essential skill during suturing and debris removal in robotic surgery, the peg transfer task entails the complex bimanual handling of small objects. Surgeons training for the daVinci (dV) robots perform peg transfers by using one or both of their arms to transfer objects between pegs [23]. Figure. 4 shows a couple of WMed surgical residents executing the peg transfer task using our Intelligent Box Trainer System (IBTS). It is required for trainees to grasp a triangle object with the grasper tool in their non-dominant hand, transfer it in mid-air to the other grasper tool in their dominant hand, and then place the object to an unoccupied peg of the pegboard to carry out the peg transfer task. A specified order by color to transfer the six triangle objects is not required. Following the transfer of all six objects to one side of the board, the trainee should then return them to the other side [24]-[27]. Ideally, the grasper tips should be moved just above the peg board and the objects should not be dropped during the transfer. In addition, there is a time limit for carrying out the test. An object detection method and a cascaded fuzzy logic supervisor were implemented to track these movements of the objects and assess the hand movements [18].

3.2.1 Object Detection Method

SSD ResNet50 V1 FPN Architecture was used to extract features for the peg transfer assessment system, and this model was trained on the created dataset. The process of tracking an object includes identifying the unique characteristics of the object to identify its unique features.

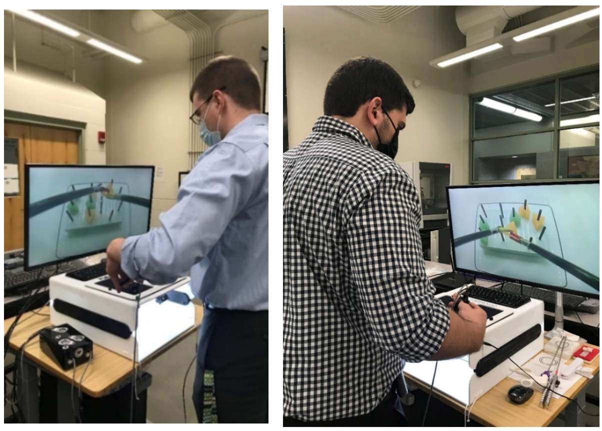
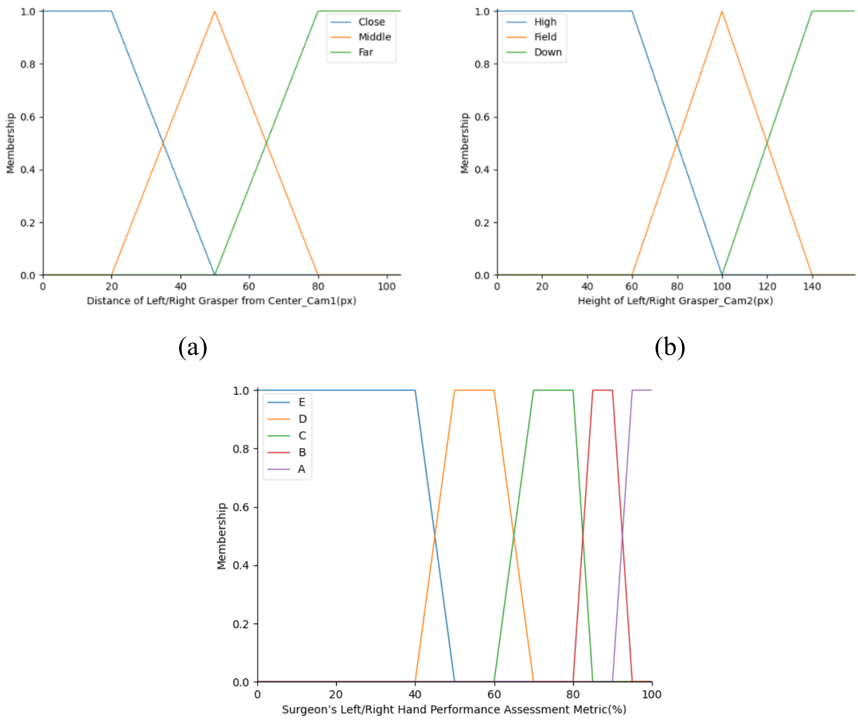


Figure 4
WMed surgical residents using the IBTS to carry out the peg transfer task

As part of this work, it is necessary to determine the related bounding box corresponding to the ground truth box to detect and track the object using the ResNet50 V1 FPN feature extractor. Considering that all the boxes are scaled differently, the best Jaccard overlap [28] should be found over a threshold greater than 0.5 to simplify the learning process. A surgeon's hand movements are monitored based on the detection and tracking of laparoscopic instruments as well as the distance between the instruments and the center of the pegboard. TensorFlow's Object Detection API was used to detect and track objects and determine their bounding boxes' coordinates. The tracking-based detection was performed frame by frame in laparoscopic video recordings. Within the IBTS, there are two side cameras and one above the pegboard, which allows the assessment to be conducted based on the results of the tracking process. A cascaded fuzzy logic supervisor is implemented to assess the hand movements by using the outcomes of the tracking method as inputs.

3.2.2 Cascaded Fuzzy Logic Supervisor

Using fuzzy logic, measured data and expert opinions can be treated in the same mathematical framework. The development of the fuzzy logic-based decision support system includes creating a knowledge base, defining an inference engine, and selecting fuzzification and defuzzification algorithms. There is a significant difference in choice depending on the application field, such as human performance assessment versus non-linear process control. Various input and output membership functions are illustrated in Fig. 5.



(a) (b)

(c)

Figure 5

(a) and (b) MFs of the input variable for the fuzzy logic evaluation system for both the right and left graspers (c) Output variable MFs for surgeon's performance assessment for the right and left-hand movements

In view of this symmetric calculation, the inputs for right-hand and left-hand movement evaluation systems are the same: the membership function (MF) and the linguistic variables (LG) of the first level, two-input MISO fuzzy logic systems working in parallel are the same for both hands. We monitored the height from the center of the pegboard in the side view scene, hence, MF and LG are different. Based on the input and output sets in the first stage of the fuzzification process, Table 1 shows the MF and LG definitions. According to Table 2, the MF and LG are shown along with their rules' values to summarize the results of both hand movements.

Table 1
Memberships and Linguistic variables for the first level fuzzy assessment

Distances	**Close** (0-50px)	**Middle** (20-80 px)	**Far** (50- ∞ px)
High (0-100 px)	B (Very Good) (80-95)	E (Fail) (0 - 50)	E (Fail) (0 - 50)
Field (60-140 px)	A (Excellent) (90-100)	B (Very Good) (80-95)	C (Good) (60 - 85)
Down (100- ∞ px)	B (Very Good) (80-95)	B (Very Good) (80-95)	D (Poor) (40 - 70)

Table 2
Memberships and linguistic variables for the second level fuzzy assessment

Left/Right Summarized Assessment	**Excellent** (50-100)	**Good** (40-60)	**Bad** (0-50)
Excellent (50-100)	A (Excellent) (90-100)	B (Very Good) (80-95)	C (Good) (60 - 85)
Good (40-60)	B (Very Good) (80-95)	C (Good) (60 - 85)	E (Fail) (0 - 50)
Bad (0-50)	C (Good) (60 - 85)	D (Poor) (40 - 70)	D (Poor) (40 - 70)

3.3 Intracorporeal Suturing

In FLS, intracorporeal suturing is part of an extensive curriculum designed to provide surgeons with an opportunity to learn and assess fundamental laparoscopic skills. There are several key components involved in intracorporeal suturing [29], [30]. For Needle Handling: the surgical instruments must be used to handle the needle accurately during laparoscopic surgery. Suturing Techniques is about suturing a set of points within a defined area (i.e., black dots on sample tissue), simulating the need for precision and dexterity during laparoscopic surgery. For Knot Tying during intracorporeal suturing, surgeons are required to demonstrate proficiency in laparoscopic knot tying. For Instrument Coordination: to achieve accurate suturing, coordination between the laparoscopic instruments and precise movements are required. For Spatial Awareness: in laparoscopic surgery, where the surgeon's view is typically two-dimensional, participants must demonstrate spatial awareness and depth perception. For Tissue Handling: particular emphasis is placed on the proper handling of tissue during suturing. Finally, for Time Management: as part of the assessment, the surgeon may be timed to determine whether he or she can perform intracorporeal suturing efficiently. Laparoscopic surgery relies heavily on time management.

3.3.1 Methods

To develop an automated assessment process, we worked with two deep-learning-based tracking algorithms, which were implemented to track the laparoscopic surgical tools in our IBTS [50]. Four metrics were used to classify and locate the surgical tool tips for assessing the trainee's performance, i.e., operation time, upward force, idle operation, and needle movement trajectory. So DL algorithms (Scaled-YOLOv4 [31], YOLOR [32], Centroid tracking [15], and DeepSORT algorithms [33]) were applied to detect and track the surgical instruments and their trajectory to measure and assess the trainee's performance. Figs. 6 and 7 illustrate the DL algorithms that were used for generating the assessment performance report and the 3D trajectory of the surgical instruments' movements.

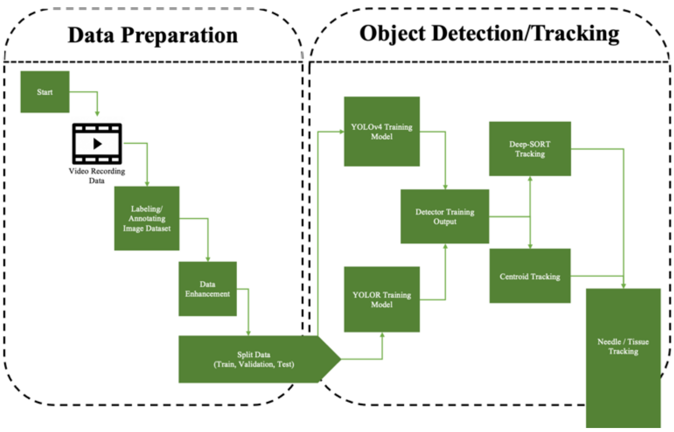


Figure 6

Proposed DL algorithms to detect and track surgical instruments for the Intracorporeal Suturing

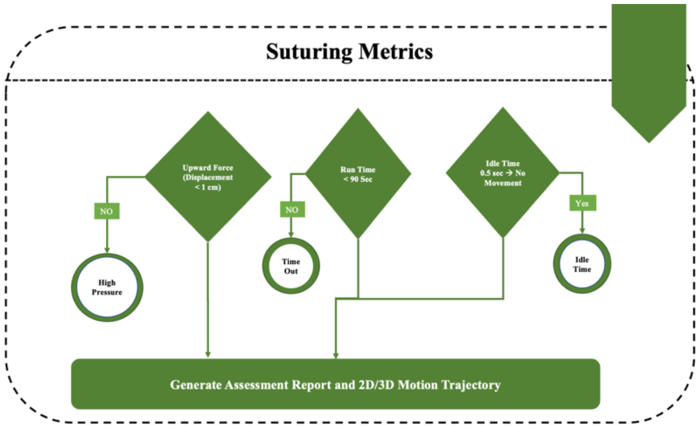


Figure 7
Inputs to the DL to generate an assessment report and movement trajectory

4 Datasets

Our automated laparoscopic surgery assessment system was developed by using the IBTS. We collected video recordings from exercises carried out by expert surgeons and novices, to extract these recordings into image frames that will feed our developed DL algorithm and fuzzy assessment system. For each of the three FLS exercises, data has been collected. As shown in Table 4, Dataset names, year, data contents, and annotation types of data are provided for each dataset specified for each FLS task. Furthermore, we included other publicly available datasets that may be of use for further research. There are several publicly available datasets, however, Table 4 includes only those datasets that are intended for the training of DL modules for FLS box trainers. As part of this paper, we intend to release three datasets for each of the FLS tasks covered here. There are 16 categories of surgical actions in SurgicalActions160, all of which are subject to surgical errors, making this dataset ideal for use as a testing resource. There are 300 videos in the LapSig300 dataset which have been collected over more than ten years, making it a rich resource for training. To support research in the Laparoscopic Surgery Skill Assessment area, our Dataset has been created and made available, at: https://drive.google.com/drive/folders/1F97CvN3GnLj-rqg1tk2rHu8x0J740DpC. As our research progresses, more files will be added to this Dataset.

Table 4
Our Proposed Datasets and Some Publicly released datasets

Name	Year	Data	Annotations
(WMU Laparoscopic Box-Trainer Dataset) Peg transfer [1]	2023	65 Videos	Tool Bounding Boxes
(WMU Laparoscopic Box-Trainer Dataset) pattern cutting [1]	2023	14 Videos	Tool Bounding Boxes
(WMU Laparoscopic Box-Trainer Dataset) suturing [1]	2023	45 Videos	Tool segmentation + Bounding Boxes
M2CAI16 [2] [34]	2016	103 videos	Gestures Labels
FlapNet [3] [35]	2020	62 minutes video	Tools segmentation
SurgicalActions160 [4] [36]	2017	160 videos	Surgical action
LapSig300 [37]	2020	300 Videos	Phases, Actions, and Tools

5 Results and Discussion

In this overview, we described the various methods we proposed and the results that we achieved. The novelty of our proposed work can be summarized as follows:

- Development of a 3D automated measurement system that monitors the surgeon's hand movements during training exercises and assesses the performance, rather than performing a visual assessment by the supervisor medical personnel.
- A combination of tracking algorithms developed to suit the nature of small objects, such as needles.
- Development of a fuzzy logic-based supervisor and assessment system to deal with the uncertainty resulting from the 2D nature of the surgical camera scene and the subjective nature of the decisions made by medical personnel.

The results of our proposed methods for each FLS exercise are illustrated in Figs. 8-10. In addition, the results of the overall mean average precision (mAP) from several DL models applied to the surgical tasks are presented in Table 5. There were some articles that used several models where other mAPs were achieved; we included the highest mAP in Table 5. As shown in Table 5, the first five references are related to our proposed work where DL algorithms were found to have a superior detection performance to other DL algorithms used in related studies. The performance assessment of a system is improved when the detection performance is higher [38].

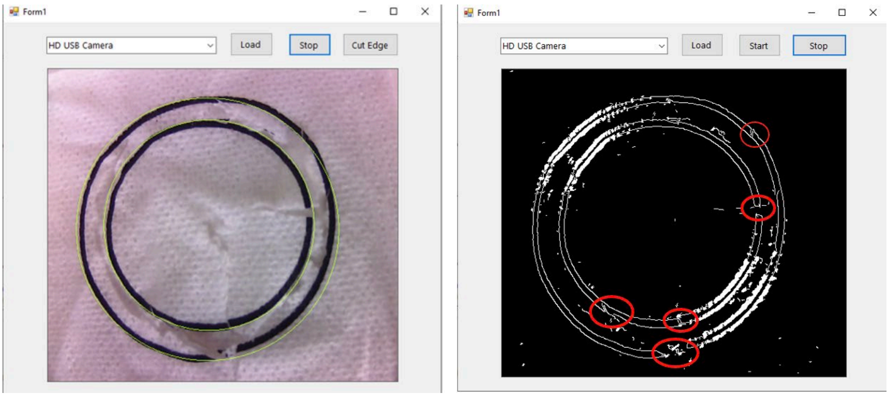


Figure 8
Pattern Cutting: shows clearly and sharply the cuts in the circle lines by applying the Canny edge detector on the captured frame

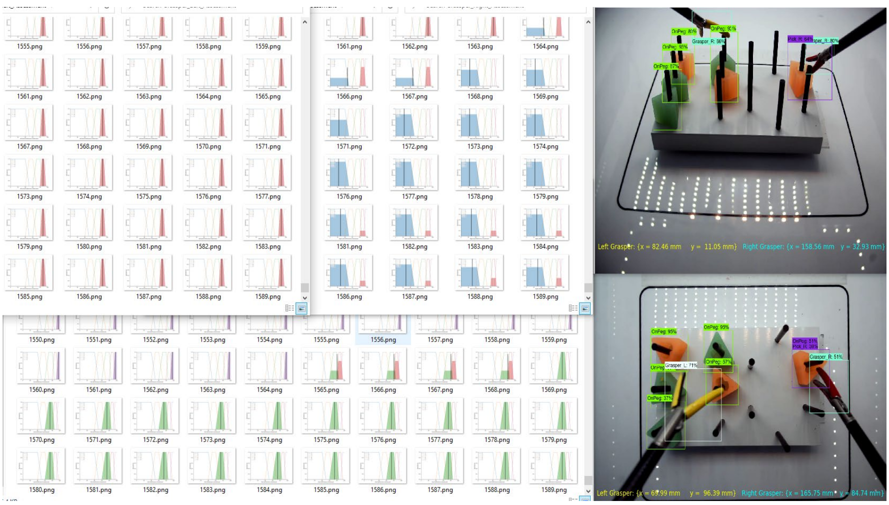


Figure 9
Peg Transfer: Object detection and measurement metric results by the top and front cameras Left/Right Hand performance assessment, and final assessment FPA, in a time period

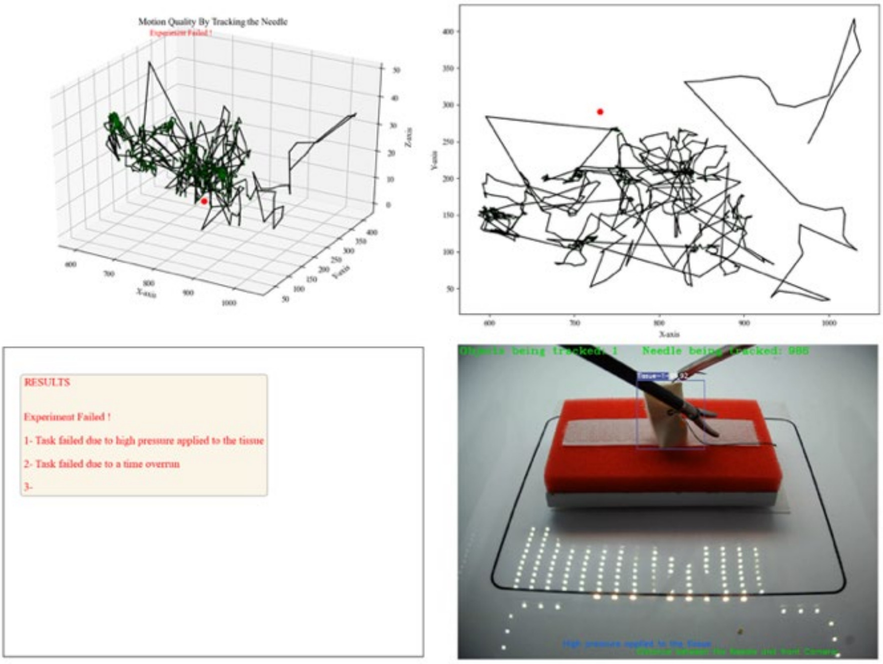


Figure 10

Intracorporeal Suturing: output of the novice surgeon using DeepSORT

Table 5

Comparison of Multi-Deep Learning algorithms applied to surgical exercises

Ref.	Year	Procedure	DL Models	Input data	Results (mAP)
[7]	2023	Suturing	DeepSORT+ YOLOR	Images	97.6%
[39]	2023	Suturing	Hybrid YOLOv8	Images	95%
[24]	2023	Peg Transfer	TensorFlow + ResNet50	Images	85%
[40]	2023	Precision Cutting	Canny Edge Detector + Blob Detector	Images	N/A
[41]	2023	Precision Cutting	YOLOv7	Images	95.1%
[42]	2020	Live Suturing	AlexNet + LSTM	Images	81%
[43]	2020	9 types of surgeries	VGG16 + LSTM	Images	75%
[44]	2019	Lab Experiments	3D CNN	Images	84.1%
[45]	2020	Lab Experiments	Deep RL	Images	81.7%
[46]	2020	Lab Experiments	VGG16	Images + Kinematic Data	86.3%

Our proposed works demonstrate the advantages of the automated assessment system. These advantages include the automatic declaration of bad performance if mistakes were made, automatic assessment for the upward forces which are applied to suturing tissue (which would not have been noticed by supervised surgeons without our system), and the tracking and recording of surgical events during the FLS task. Furthermore, by using an automated assessment like the one proposed in our proposal, FLS tests can be practiced without the supervision by an expert surgeon. According to all proposed works, recording videos yielded better results than real-time experiments. This is due to the lack of suitable computing power. The algorithms used in all proposed works resulted in a delay of 10-15 seconds in the detection process. Our objective is to develop a more powerful IBTS by installing a high-performance GPU that can handle real-time experiments better.

Concluding Remarks

The purpose of this paper is to overview recent research projects in the Fuzzy/Intelligent Control Systems Laboratory at WMU to develop an automated skill assessment system for FLS tests using the IBTS. A system based on an automated assessment system can avoid potential problems associated with incorrect decisions made by supervisor medical personnel (e.g., misidentification of situations). In addition, the recorded test videos along with the assessments can facilitate and accelerate the trainees' efforts to improve their performance. Additional contributions of this study are related to providing tracking and assessing the hand movements in 3D space and upward forces in case of suturing, Throughout our proposed work, DL models (including Hybrid YOLOv8, YOLOR, YOLOv7, TensorFlow Object Detection API, and DeepSORT) and fuzzy logic decision support systems were discussed and implemented. Several experiments demonstrated that the proposed algorithms performed better in recorded assessment videos than in real-time experiments due to the 10-15 second delay in detecting objects. To improve the real-time performance of the system, the IBTS will need a powerful GPU. Over the last 6 years, our datasets have been developed, such that our extracted video images are labeled by using inputs from expert surgeons. To facilitate research in the area of laparoscopic surgery skill assessment, we have created a dataset and made it publicly available. In the future, we also plan to investigate the use of intelligent robotic arms and hands in laparoscopic surgery procedures and assessment.

Reference

[1] W. Reynolds, "The first laparoscopic cholecystectomy.," *JSLS*, Vol. 5, No. 1, pp. 89-94, 2001

[2] B. Jaffray, "Minimally invasive surgery," *Arch. Dis. Child.*, Vol. 90, No. 5, pp. 537-542, 2005, doi: 10.1136/adc.2004.062760

[3] A. Goldbraikh, A. L. D'Angelo, C. M. Pugh, and S. Laufer, "Video-based fully automatic assessment of open surgery suturing skills," *Int. J. Comput. Assist. Radiol. Surg.*, Vol. 17, No. 3, pp. 437-448, 2022, doi:

10.1007/s11548-022-02559-6

[4] A. Chellali *et al.*, “Achieving interface and environment fidelity in the Virtual Basic Laparoscopic Surgical Trainer,” *Int. J. Hum. Comput. Stud.*, Vol. 96, pp. 22-37, 2016

[5] J. H. Peters *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

[6] A. G. Gallagher *et al.*, “Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training,” *Ann. Surg.*, Vol. 241, No. 2, pp. 364-372, 2005, doi: 10.1097/01.sla.0000151982.85062.80

[7] M. Mohaidat, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, “Instrument Detection for the Intracorporeal Suturing Task in the Laparoscopic Box Trainer Using Single-stage object detectors,” *IEEE Int. Conf. Electro Inf. Technol.*, Vol. 2022-May, pp. 455-460, 2022, doi: 10.1109/eIT53891.2022.9813888

[8] H. M. Oh and M. Y. Kim, “Attitude tracking using an integrated inertial and optical navigation system for hand-held surgical instruments,” *Int. Conf. Control. Autom. Syst.*, Vol. 17, No. 2, pp. 290-293, 2014, doi: 10.1109/ICCAS.2014.6988005

[9] I. Oropesa *et al.*, “Methods and Tools for Objective Assessment of Psychomotor Skills in Laparoscopic Surgery,” *J. Surg. Res.*, Vol. 171, No. 1, pp. e81-e95, 2011, doi: https://doi.org/10.1016/j.jss.2011.06.034

[10] J. Lynch, P. Aughwane, and T. M. Hammond, “Video games and surgical ability: a literature review.,” *J. Surg. Educ.*, Vol. 67, No. 3, pp. 184-189, 2010, doi: 10.1016/j.jsurg.2010.02.010

[11] A. Shrestha and A. Mahmood, “Review of deep learning algorithms and architectures,” *IEEE access*, Vol. 7, pp. 53040-53065, 2019

[12] I. Rivas-Blanco, C. J. Pérez-Del-Pulgar, I. García-Morales, and V. F. Muñoz, “A review on deep learning in minimally invasive surgery,” *IEEE Access*, Vol. 9, pp. 48658-48678, 2021

[13] M. Mohaidat, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, “A Hybrid YOLOv8 and Instance Segmentation Approach Towards Autonomous Suturing Performance Assessment,” in *2023 IEEE 5th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE)*, Oct. 2023, p. Accepted

[14] F. R. Fathabadi, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, “Surgical Skill Assessment System Using Fuzzy Logic in a Multi-Class Detection of Laparoscopic Box-Trainer Instruments,” in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2021, pp. 1248-1253,

doi: 10.1109/SMC52423.2021.9658766

[15] M. Mohaidat, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, "Multi-class Detection and Tracking of Intracorporeal Suturing Instruments in an FLS Laparoscopic Box Trainer Using Scaled-YOLOv4," in *Advances in Visual Computing. ISVC 2022*, G. Bebis, Ed. Springer International Publishing, 2022, pp. 211-221. doi: 10.1007/978-3-031-20713-6_16

[16] F. R. Fathabadi, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, "Surgical Skill Training and Evaluation for a Peg Transfer Task of a Three Camera-Based Laparoscopic Box-Trainer System," pp. 1146-1151, 2022, doi: 10.1109/csci54926.2021.00242

[17] K. N. Alkhamaiseh, J. L. Grantner, S. Shebrain, and I. Abdel-Qader, "Towards Automated Performance Assessment for Laparoscopic Box Trainer using Cross-Stage Partial Network"

[18] F. Rashidi Fathabadi, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, "3D Autonomous Surgeon's Hand Movement Assessment Using a Cascaded Fuzzy Supervisor in Multi-Thread Video Processing," *Sensors*, Vol. 23, No. 5, 2023, doi: 10.3390/s23052623

[19] M. Ilyas, M. F. U. Butt, M. Bilal, K. Mahmood, A. Khaqan, and R. Ali Riaz, "A Review of Modern Control Strategies for Clinical Evaluation of Propofol Anesthesia Administration Employing Hypnosis Level Regulation," *Biomed Res. Int.*, Vol. 2017, p. 7432310, 2017, doi: 10.1155/2017/7432310

[20] A. Rikalovic and I. Cosic, "A fuzzy expert system for industrial location factor analysis," *Acta Polytech. Hungarica*, Vol. 12, No. 2, pp. 33-51, 2015

[21] H. Kong, H. C. Akakin, and S. E. Sarma, "A generalized Laplacian of Gaussian filter for blob detection and its applications," *IEEE Trans. Cybern.*, Vol. 43, No. 6, pp. 1719-1733, 2013

[22] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, No. 6, pp. 679-698, 1986

[23] G. T. Gonzalez *et al.*, "From the dexterous surgical skill to the battlefield—a robotics exploratory study," *Mil. Med.*, Vol. 186, No. Supplement_1, pp. 288-294, 2021

[24] F. Rashidi Fathabadi, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, "Autonomous sequential surgical skills assessment for the peg transfer task in a laparoscopic box-trainer system with three cameras," *Robotica*, pp. 1-19, 2023, doi: 10.1017/s0263574723000218

[25] J. Zhang and X. Gao, "Object extraction via deep learning-based marker-free tracking framework of surgical instruments for laparoscope-holder robots," *Int. J. Comput. Assist. Radiol. Surg.*, Vol. 15, pp. 1335-1345, 2020

[26] K. Kunkler, "The role of medical simulation: an overview," *Int. J. Med. Robot. Comput. Assist. Surg.*, Vol. 2, No. 3, pp. 203-210, 2006

[27] J. L. Grantner, A. H. Kurdi, M. Al-Gailani, I. Abdel-Qader, R. G. Sawyer, and S. Shebrain, "Intelligent Performance Assessment System for Laparoscopic Surgical Box-Trainer," in *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2018, pp. 1-7

[28] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, "Scalable object detection using deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 2147-2154

[29] P. Crochet, A. Agostini, S. Knight, N. Resseguier, S. Berdah, and R. Aggarwal, "The Performance Gap for Residents in Transfer of Intracorporeal Suturing Skills From Box Trainer to Operating Room," *J. Surg. Educ.*, Vol. 74, No. 6, pp. 1019-1027, 2017, doi: https://doi.org/10.1016/j.jsurg.2017.05.013

[30] D. Qi, K. Panneerselvam, W. Ahn, V. Arikatla, A. Enquobahrie, and S. De, "Virtual interactive suturing for the Fundamentals of Laparoscopic Surgery (FLS)," *J. Biomed. Inform.*, Vol. 75, pp. 48-62, Nov. 2017, doi: 10.1016/j.jbi.2017.09.010

[31] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Scaled-yolov4: Scaling cross stage partial network," in *Proceedings of the IEEE/cvf conference on computer vision and pattern recognition*, 2021, pp. 13029-13038

[32] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "You Only Learn One Representation: Unified Network for Multiple Tasks," May 2021 [Online] Available: http://arxiv.org/abs/2105.04206

[33] L. He, G. Liu, G. Tian, J. Zhang, and Z. Ji, "Efficient Multi-View Multi-Target Tracking Using a Distributed Camera Network," *IEEE Sens. J.*, Vol. 20, No. 4, pp. 2056-2063, 2020, doi: 10.1109/JSEN.2019.2949385

[34] A. Jin *et al.*, "Tool detection and operative skill assessment in surgical videos using region-based convolutional neural networks," in *2018 IEEE winter conference on applications of computer vision (WACV)*, 2018, pp. 691-699

[35] A. Attanasio *et al.*, "Autonomous Tissue Retraction in Robotic Assisted Minimally Invasive Surgery – A Feasibility Study," *IEEE Robot. Autom. Lett.*, Vol. 5, No. 4, pp. 6528-6535, 2020, doi: 10.1109/LRA.2020.3013914

[36] K. Schoeffmann, H. Husslein, S. Kletz, S. Petscharnig, B. Muenzer, and C. Beecks, "Video retrieval in laparoscopic video recordings with dynamic content descriptors," *Multimed. Tools Appl.*, Vol. 77, No. 13, pp. 16813-16832, 2018, doi: 10.1007/s11042-017-5252-2

[37] D. Kitaguchi *et al.*, "Automated laparoscopic colorectal surgery workflow recognition using artificial intelligence: experimental research," *Int. J. Surg.*, Vol. 79, pp. 88-94, 2020

[38] F. R. Fathabadi, J. L. Grantner, I. Abdel-Qader, and S. A. Shebrain, "Box-Trainer Assessment System with Real-Time Multi-Class Detection and

Tracking of Laparoscopic Instruments, using CNN," *Acta Polytech. Hungarica*, Vol. 19, No. 2, 2022

[39] M. Mohaidat, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, "A Hybrid YOLOv8 and Instance Segmentation to Distinguish Sealed Tissue and Detect Tools' Tips in FLS laparoscopic box trainer," in *2023 IEEE SMC Conference*, p. 2-page Abstract

[40] K. N. Alkhamaiseh, J. L. Grantner, I. Abdel–Qader, and S. Shebrain, "A Summative Assessment of the Pattern-Cutting Task in the Laparoscopic Box Trainer using Color Tracking and Fuzzy Logic," *Jordan J. Electr. Eng.*, 2023

[41] K. N. Alkhamaiseh, J. L. Grantner, I. Abdel-qader, and S. Shebrain, "Towards Real-Time Multi-Class Object Detection and Tracking for the FLS Pattern Cutting Task," *Adv. Sci. Technol. Eng. Syst. J.*, Vol. 95, No. 6, pp. 87-95, 2023, doi: 10.25046/aj080610

[42] F. Luongo, R. Hakim, J. H. Nguyen, A. Anandkumar, and A. J. Hung, "Deep learning-based computer vision to recognize and classify suturing gestures in robot-assisted surgery," *Surgery*, Vol. 169, No. 5, pp. 1240-1244, 2021

[43] S. Kannan, G. Yengera, D. Mutter, J. Marescaux, and N. Padoy, "Future-state predicting LSTM for early surgery type recognition," *IEEE Trans. Med. Imaging*, Vol. 39, No. 3, pp. 556-566, 2019

[44] I. Funke, S. Bodenstedt, F. Oehme, F. von Bechtolsheim, J. Weitz, and S. Speidel, "Using 3d convolutional neural networks to learn spatiotemporal features for automatic surgical gesture recognition in video," in *International conference on medical image computing and computer-assisted intervention*, 2019, pp. 467-475

[45] X. Gao, Y. Jin, Q. Dou, and P.-A. Heng, "Automatic gesture recognition in robot-assisted surgery with reinforcement learning and tree search," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 8440-8446

[46] Y. Qin *et al.*, "Temporal segmentation of surgical sub-tasks through deep learning with multiple data sources," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 371-377