


RESEARCH ARTICLE

Autonomous sequential surgical skills assessment for the peg transfer task in a laparoscopic box-trainer system with three cameras

Fatemeh Rashidi Fathabadi^{1,*} , Janos L. Grantner¹, Saad A Shebrain² and Ikhlas Abdel-Qader¹

¹Department of Electrical & Computer Engineering, Western Michigan University, Kalamazoo, MI, USA and ²Homer Stryker M.D. School of Medicine, Western Michigan University, Kalamazoo, MI, USA

*Corresponding author. E-mail: fatemeh.rashidifathabadi@wmich.edu

Received: 21 August 2022; **Revised:** 2 January 2023; **Accepted:** 4 February 2023; **First published online:** 3 March 2023

Keywords: FLS assessment, peg transfer task, object detection, convolutional neural network, laparoscopic surgery, laparoscopic box trainer

Abstract

In laparoscopic surgery, surgeons should develop several manual laparoscopic skills before carrying out real operative procedures using a low-cost box trainer. The Fundamentals of Laparoscopic Surgery (FLS) program was developed as a program to assess fundamental knowledge and surgical skills, required for basic laparoscopic surgery. The peg transfer task is a hands-on exam in the FLS program that assists a trainee to understand the relative minimum amount of grasping force necessary to move the pegs from one place to another place without dropping them. In this paper, an autonomous, sequential assessment algorithm based on deep learning, a multi-object detection method, and, several sequential If-Then conditional statements have been developed to monitor each step of a surgeon's performance. Images from three different cameras are used to assess whether the surgeon executes the peg transfer task correctly and to display a notification on any errors on the monitor immediately. This algorithm improves the performance of a laparoscopic box-trainer system using top, side, and front cameras and removes the need for any human monitoring during a peg transfer task. The developed algorithm can detect each object and its status during a peg transfer task and notifies the resident about the correct or failed outcome. In addition, this system can correctly determine the peg transfer execution time, and the move, carry, and dropped states for each object by the top, side, and front-mounted cameras. Based on the experimental results, the proposed surgical skill assessment system can identify each object at a high score of fidelity, and the train-validation total loss for the single-shot detector (SSD) ResNet50 v1 was about 0.05. Also, the mean average precision (mAP) and Intersection over Union (IoU) of this detection system were 0.741, and 0.75, respectively. This project is a collaborative research effort between the Department of Electrical and Computer Engineering and the Department of Surgery, at Western Michigan University.

1. Introduction

For certain surgical procedures, Minimally Invasive Surgery (MIS) has become more advantageous than traditional open surgery and represents a turning point in the history of laparoscopic surgery [1, 2]. Carrying out a high-quality laparoscopic surgery procedure is a challenging task for any surgeon. In addition, the surgeon's skills to perform MIS should also be enhanced through training [3, 4]. Toward that goal, several simple training tasks such as peg transfers, clip and divide, pattern cutting, suturing with intracorporeal or extracorporeal knots, and fixation have been developed [5].

The peg transfer task is one of the hands-on exams in the Fundamental Laparoscopic System (FLS) [6, 7] in which a trainee works with two graspers. In recent years, several peg transfer skill assessment systems have been proposed using laparoscopic box trainers and daVinci robots. In this regard, a variety

of computer vision applications for tooltip detection, tooltip tracking, and motion detection have been developed for laparoscopic surgery training and performance assessment [8].

By reviewing previous works, most of the research has been based on color detection, Haar wavelets, texture features, and gradient-based features, while none of them are stable methods. Color-based detection is one of the traditional ways to find the laparoscopic instruments and their coordinates on the test platform during exercises. However, this method could not create the label of each instrument. Based on the literature review, current autonomous skill assessments are divided into two categories: generic aggregate metrics and statistical models. Generic aggregate metrics assess the surgeons' skills and performance in structured tasks using box trainers and virtual reality (VR) simulators [9]. On the other hand, in the case of the laparoscopic box trainer, most motion analysis, and tooltip tracking evaluations have used statistical metrics such as speed, acceleration, total time, path length, number of movements, and some other parameters, which would not help trainees in performing the procedure correctly [10].

Moreover, in previous research, the performance evaluation metrics for FLS box trainers were not based on all the mandatory sequential steps, neither they proposed an autonomous skill assessment. Based on our study, there is no published research that classifies the level of a surgeon's performance and sequentially tracks all the procedures and evaluates performance step by step and alerts the trainees in case of errors. Besides, most of the previous works in the field of laparoscopic box trainer, based on deep learning and machine learning, did not focus on instrument detection or multi-instrument detection. They did not lead to performance tracking and autonomous skill assessment and were mostly based on skill-level classification using time series data. Although deep learning and machine learning algorithms were the basis of previous recent works, most of them addressed laparoscopic instrument detection in the field of robotic surgery. To our best knowledge, several challenges are still open in the field of performance assessment for using laparoscopic box trainer such as:

- Building a laparoscopic box-trainer custom dataset including the peg transfer task.
- Developing a robust multiclass object detection method, which classifies and recognizes each object separately, instead of the colored tooltip detection method. Dying the tooltips does not provide a durable solution, and color filtering algorithms are heavily affected by changing lighting conditions.
- Developing an autonomous sequential peg transfer task evaluation method for a laparoscopic box trainer.

The purpose of this study is to propose an autonomous laparoscopic surgery performance assessment system using a combination of artificial intelligence (AI) and a box-trainer system. The novelty of this work is an autonomous, sequential assessment of the performance of a resident or surgeon during the peg transfer task, rather than doing that by an observing supervisor personnel. In other words, to assess the performance of laparoscopic residents and surgeons, their actions have to be recorded and assessed by an expert. To prevent any subjective errors in the decisionmaking, an autonomous assessment system should be developed. Human intervention, in particular in training and evaluation of laparoscopic trainees' performance, is not only time-consuming but expensive. One of the advantages of this autonomous skill assessment is that even if a trainee or the expert that is evaluating the trainee did not notice any mistake, the system would record and report any wrong performance. Also, using the autonomous assessment approach will open more time for residents to practice without having a supervisor surgeon present there. The proposed algorithm has the capability of detecting each peg transfer step separately and notifying the resident on the outcome of the test.

The paper is organized as follows: Section 2 gives a review of related work in this research area. In Section 3, a detailed explanation of the methodology employed in this research is outlined. In Section 4, the experimental results are presented and discussed. In Section 5, conclusions and future plans are given.

1.1. Related works

1.1.1. Peg transfer task in the daVinci robot

The applications of AI in robotic surgery and autonomous robotics have been significantly increased over the last decade [11–13]. Since the year 2000, when the daVinci Surgical System [14] was approved by the United States Food and Drug Administration (FDA), the use of this first robotic platform in surgical procedures has created a new era in MIS [15]. Several surgical fields all around the world have adopted the daVinci (dV) [16]. Surgical system platform and researchers have studied various topics associated with robotic surgery such as surgical instrument segmentation [17, 18], workflow recognition [19], gesture recognition [20, 21], and surgical scene reconstruction [22]. All this research has generated a valuable database for surgical training and evaluation [23].

However, the most obvious concern in robotic surgery is related to surgical errors resulting in disability or death of the patient [24]. To address this issue, six key tasks in laparoscopic surgery training (which is a requirement for robotic surgery certification) have been developed. The peg transfer task is one of them, which is comprised of complex bimanual handling of small objects, as an essential skill during suturing and debris removal in robotic surgery [25]. In the dV robot, the peg transfer task is done with one or both of its arms to transfer the objects between pegs. Hannaford et al. [26] used a single Raven II [27] arm to perform handover-free peg transfer with three objects. Hwang [28] considered a monochrome (all red) version of the peg transfer task, implemented with the daVinci Research Kit (dVRK) for a robust, depth-sensing perception algorithm.

Hwang extended his research in ref. [29] and developed a method that integrates deep learning and depth sensing. In his work, a dVRK surgical robot with a Zivid depth sensor was used to automate three variants of a peg transfer task. 3D-printed fiducial markers with depth sensing and a deep recurrent neural network have been used to increase the accuracy of the dVRK to less than 1 mm. Based on the results, this autonomous system could outperform an experienced human surgical resident in terms of both speed and success rate. Qin et al. [30] compared simulators and HTC Vive and Samsung Gear devices to develop and assess the peg transfer training module and also to compare the advantages and disadvantages of the box trainer, the VR, AR, and MR trainers. Brown et al. [31] instrumented the dV Standard robot with their own Smart Task Board (STB) and recruited 38 participants of varying skills in robotic surgery exercises to perform three trials of peg transfer. The surgeon's skill assessment at the robotic peg transfer was based on regression and classification machine learning algorithms, using features extracted from the force, acceleration, and time sensors external to the robot.

Aschwanden and Burgess et al. [32] used a VR manual skills experiment to compare human performance measures with experiences which were indicated on a questionnaire handed out. In this experiment, among video games and computer proficiency, accuracy, time, efficiency of motion, and number of errors were the main parameters in performance measures. Based on authors' results, there was little or no relations between performance and experience level. Coad et al. [33] developed a STB consisting of an accelerometer clip for the robotic camera arm, accelerometer clips for the two primary robotic arms, a task platform containing a three-axis force sensor, a custom signal conditioning box, and an Intel NUC computer for data acquisition to perform a peg transfer task. In addition, a proportional-derivative (PD) controller was developed to control the position and orientation of the master gripper. In this research, the physical interaction data from the patient-side manipulators of a dV surgical system were recorded to consider the effects of null, divergent, and convergent force/torque in the peg transfer task.

Chen et al. [34] developed a haptic rendering interface based on Bayesian optimization to tune user-specific parameters during the surgical training process. This research aimed to assess the differences between with and without using the supervised semi-autonomous control method for a surgical robot and was validated in a customized simulator based on a peg transfer task. Hutchins et al. in ref. [35] recruited a total of 21 right-hand surgeons with varying surgical skills to perform five iterations of the peg transfer task from the FLS manual skills curriculum using two Maryland forceps. In this study, the smoothness of the sporadic movement of the Maryland forceps was evaluated with the speed power spectrum based on capturing the trace trajectories using an electromagnetic motion tracking system.

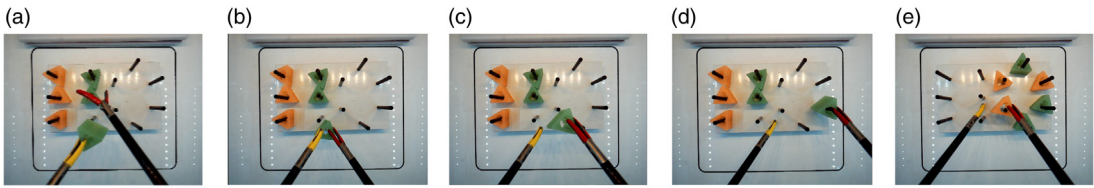


Figure 1. Pegboard with 12 pegs, and six rubber ring-like objects (triangles) in a three-camera laparoscopic box-trainer system.

Zhao and Voros et al. in ref. [36] proposed a 2D/3D tracking-by-detection framework which could detect shaft and end-effector portions. The shaft was detected by line features via the RANdom Sample Consensus (RANSAC) scheme, in which finds an object pose from 3D–2D point correspondences to fit lines passing through a set of edge points, while the end effector was detected by image features based on a modified AlexNet model. Zhao and Voros have also developed a real-time method for detection and tracking of two-dimensional surgical instruments in ref. [37] based on a spatial transformer network (STN) for tool detection and spatio-temporal context (STC) for tool tracking.

1.1.2. Peg transfer task in a laparoscopic box trainer

To prepare a trainee for a real operation, several laparoscopic box-trainer skill tests should be mastered [38]. The peg transfer task is the first hands-on exam in the FLS in which a trainee should grasp each triangle object with a non-dominant hand Fig. 1(a) and transfer it in mid-air to the dominant hand, as illustrated in Fig. 1(b) and (c). Then the triangle object should be put on a peg on the opposite side of the pegboard as shown in Fig. 1(d). There is no requirement on the color or order in which the six triangle objects are transferred. Once all six objects have been transferred to the opposite side of the board as shown in Fig. 1(e), the trainee should carry out transferring them back to the left side of the board [39–41].

Gao et al. [42] analyzed the learning curve data to train a multivariate, supervised machine learning model known as the kernel partial least squares, in a pattern-cutting task on an FLS box trainer and a virtual basic laparoscopic skill trainer (VBLaST) over 3 weeks. Also, an unsupervised machine learning model was utilized for trainee classification. Using machine learning algorithms, the authors assessed the performance of trainees by evaluating their performance during the first 10 repetitive trials and predicted the number of trials required to achieve proficiency and final performance level. However, there is lack of information regarding how the dataset was prepared, what the evaluation criteria was and there was no visualization.

Kuo et al. [43] developed a method for skill assessment the peg transfer task in a laparoscopic simulation trainer based on machine learning algorithms to distinguish between experts and novices. In this work, two types of datasets, including time series data extracted from the coordinates of eye fixation and image data from videos, were used. The movement of the eyes was recorded as a time series of gaze point coordinates, and it was the input to a computer program to find a classification of the surgeons' skills. Also, images were extracted from videos, where they were labeled according to the corresponding skill levels and they were fed into the Inception-ResNet-v2 to classify the image data. Finally, the outcomes from the time series and the image data were used as inputs to extreme learning machines to assess a trainee's performance.

A fuzzy reasoning system was designed by Hong et al. [44] to assist with visual guidance and the assessment of trainees' performance for the peg transfer task. The fuzzy reasoning system was made up of five subsystems such as the distance between an instrument's tip position and a target, the state of the grasper to check if the grasper was open or not, a support vector machine algorithm to classify whether an object was on a peg or out of it, dimensions of the triangle-shaped object to estimate if the object was on the peg or was close to the camera, and finally, the distance between an instrument's tip position and

the center of the object. The fuzzy rules were designed based on the above-mentioned inputs to assess the performance of a trainee.

Islam et al. [45] developed a web-based tool in which surgeons can upload videos of their MIS training and assess their performance by using the presented evaluation scores and analysis over time. This methodology was based on object segmentation, motion detection, feature extraction, and error detection. Their algorithm detected triangular objects in the peg transfer task and computed the total number of objects by counting all the pixels inside the bounding box. The errors were recorded based on the number of objects on the pegs. If the objects were not inside the bounding box, then the missing number of objects was recorded as errors. Peng et al. [46] developed a Single-Shot State Detection (SSSD) method using deep neural networks for the peg transfer task on 2D images. In this study, a deep learning method and then the single-shot detector (SSD) was used to detect the semantic objects, and then the IoU was measured as the assessment metric of localization.

Jiang et al. [47] proposed a laparoscopic trainer system by adding a 3D augmented reality (AR) visualization to it for the classic peg transfer task. The authors predicted the pose of the triangular prism from a single RGB image by training a Pose CNN architecture that could track the laparoscopic grasper. Meisner et al. [48] designed an object state estimator and a tracking method for the peg transfer task. This method supported visual and force guidance for a computer-assisted surgical trainer (CAST) using image processing approaches. The object state estimator included three subsystems such as color segmentation in a Hue–Saturation–Value (HSV) color domain, object tracking using Harris Corner Detection, and a shape approximation algorithm [49]. Peng et al. [50] designed a simulation process for an image-based object state modeling technique to develop an object state tracking system for the peg transfer task. A rule-based, intelligent method had been utilized to discriminate the object state without the support of any object template.

Pérez-Escamirosa et al. [51] developed a sensor-free system to detect and track movements of laparoscopic instruments in a 3D workspace based on green and blue color markers which were placed on the tip of the instruments. In ref. [52], researchers proposed a tracking system using a sequence of images, color segmentation, and the Sobel Filtering method. In this method, information obtained from the laparoscopic instrument's shaft edge was extracted to detect motion fields of the laparoscopic instruments via video tracking. Allen et al. [53] predicted a tooltip position of a laparoscopic instrument based on the instrument's shaft detection. In this research, a color space analysis was applied to extract the instrument contours, as well as fitting a line to estimate the direction of each laparoscopic instrument's movement. Finally, a linear search was employed to identify the position of each instrument's tooltip.

Two elements, eye-hand coordination and manual dexterity, are critical for a successful, uneventful procedure and for the patient's safety during laparoscopic surgery. In laparoscopic surgery, the contact between the surgeon's hands and tissues or organs of the patient is maintained indirectly through the laparoscopic instruments [54]. In addition, the visualization of the surgical field and the steps of the procedure carried out on the tissues or organs is implemented on a computer monitor through a microcomputer system. Despite using the laparoscopic instruments, haptics (transmitting information through the sense of touch) are relatively preserved [55]. To adapt to this new interface, practice is essential. Therefore, using peg transfer, the trainees can understand the relative minimum amount of grasping force necessary to move the pegs from one place to another without dropping them.

Chen et al. [56] reviewed 66 studies in the field of laparoscopy in which the majority of studies had focused on the assessment of a trainee's performance in basic tasks such as suturing, peg transfer, and knot tying. Furthermore, the authors mentioned that most of these studies were derived from the daVinci robot, and 10 studies used a simulator with external sensors attached on the instruments such that the surgeons were able to track instruments' movements. Most of the studies measured surgical performance based on the classification of trainees' skills into novices or experts using machine learning algorithms and cross-validation techniques (a method for assessing the classification ability of the machine learning model) or a simple neural network model. However, other studies intended to predict skill-level scores on global rating scales. In our study, we assess the trainee's performance by sequentially monitoring

the actions and the hands' movements and giving the feedback based on several skill assessment criteria which are all based on multi-object detection. Moreover, reviewing all these published studies, the authors stated that the available datasets are small. In addition, the majority of data obtained from these studies are not open source or only related to robotic surgery, not a box trainer or a simulator. In our study, we share our dataset with other researchers who are interested in FLS box-trainer-related projects, and our code will be posted after the paper is published.

2. Methodology

This activity did not constitute human subjects research based on the definition provided by the Common Rule; therefore, documented informed consent was not required. However, verbal consent to participate was obtained.

To carry out the peg transfer task successfully, the trainee has to perform all of steps outlined before. If the number of dropped objects exceeds two in the field of interest, or if an object falls out of the field, or the trainee could not finish the full peg transfer task within the maximum time allowed for the whole task execution, then the test is failed. To monitor and then assess all the moments of a trainee's performance, a multi-object detection algorithm [57] based on a deep neural network architecture [58] is proposed which records the trainee's performance by using top, side, and front cameras. These cameras monitor all the steps performed by the trainee from three different perspectives and provide the inputs for the multi-object detection [59] and sequential assessment algorithms. An overview of the proposed algorithm for the sequential assessment and surgical skills training system for the peg transfer task using the three camera based method is depicted in Fig. 2.

The sequential surgical skill assessment block, shown in Figs. 5 and 6, respectively, is a consecutive algorithm. It is executed in a sequential manner, meaning that when the trainee starts performing the peg transfer task, the algorithm is carried out step by step without any other processing is executed. The sequential surgical skill assessment algorithm has been developed based on several sequential If-Then conditional statements to monitor each step of the surgeon's actions sequentially with three cameras assess whether the surgeon is performing the peg transfer task in the correct order and display a notification on the monitor immediately.

2.1. Data preparation

The first aim of this research is to propose a deep learning architecture for detecting and tracking all the intended objects, as listed in Table I, during the peg transfer task in the field of interest. To accomplish that, a large number of laparoscopic box-trainer images, related to the peg transfer task, need to be acquired from the three cameras of our laparoscopic box trainer.

2.2. Experimental setup

Our Intelligent Box-Trainer System (IBTS) has been technologically advanced for laparoscopic surgery skill assessment studies. The main components of the IBTS system used for this experiment are the 32" HD computer monitor and three 5-megapixel USB 2.0 cameras with variable focus lens which are mounted on top, on side wall, and in front of the box [60]. The system has more components that are dedicated to pursuing other research objectives as well. The three cameras have been adjusted to make sure that the field of view is centered on the pegboard, and the entire pegboard along with a marker line indicating the field of interest are visible. The IBTS also has the capability to measure the forces applied by the grasper jaws. In case of actual laparoscopic surgery, the visual cues learned during experiments are important when manipulating the tissues, such as the intestines, where extreme caution is needed to avoid using unnecessary high grasping forces that can cause damage [61].

The sequential surgical skills assessment algorithm for the peg transfer task is one of the enhancements. It has been implemented on our IBTS by using several sequential If-Then conditional statements,

Table I. Defined labels for each object in each frame.

Instrument	Object activity definition table
OnPeg	A triangle object is not connected to the grasper but is on a peg
OutPeg	A triangle object is not connected to the grasper and is not on any peg or in the field
Grasper_R	The right grasper
Grasper_L	The left grasper
Carry_R	A triangle is carried by the right grasper but is not on any peg
Carry_l	A triangle is carried by the left grasper but is not on any peg
Transfer	A triangle is being transferred from one grasper to another grasper and is not on any peg
Pick_R	A triangle is picked by the right grasper and is on a peg or in the field
Pick_L	A triangle is picked by the left grasper and is on a peg or in the field

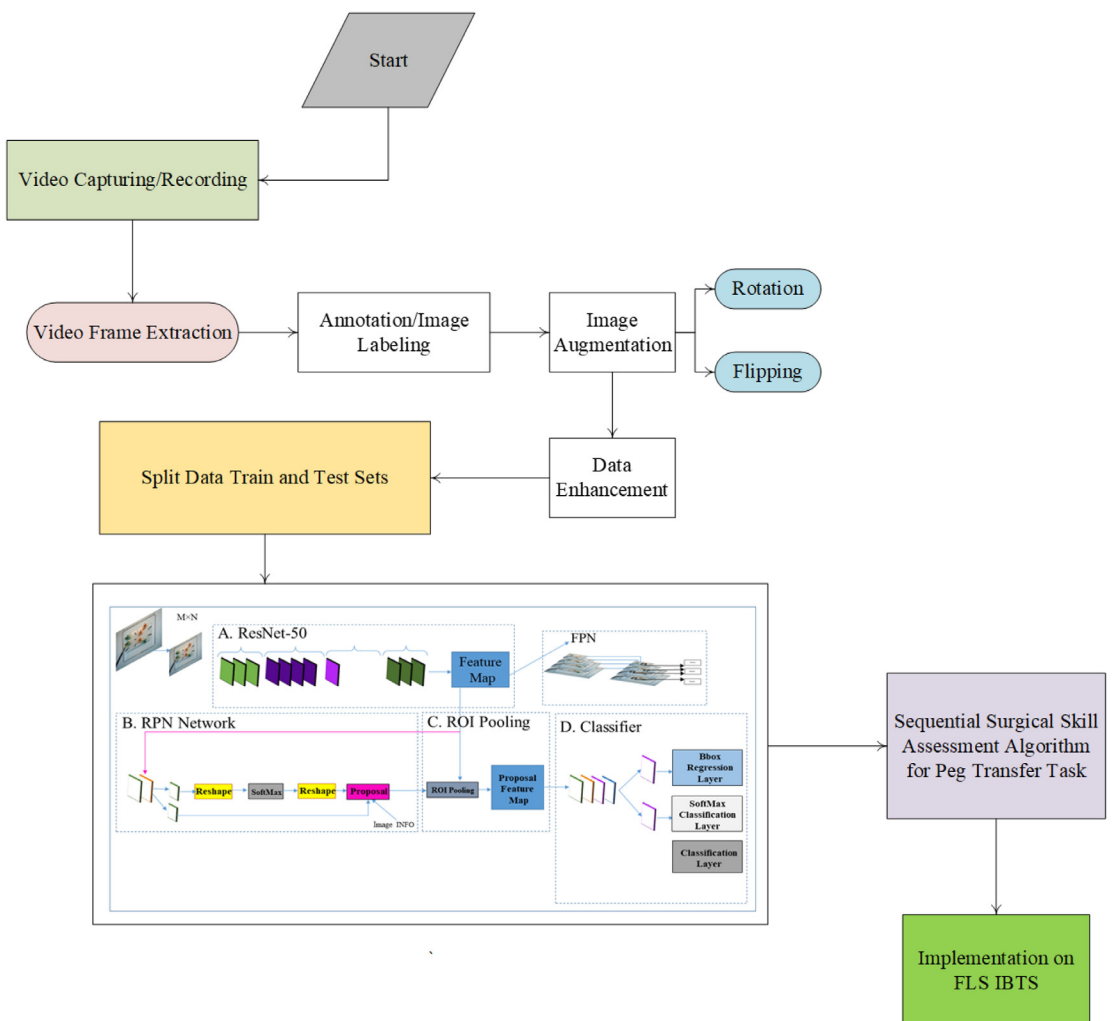
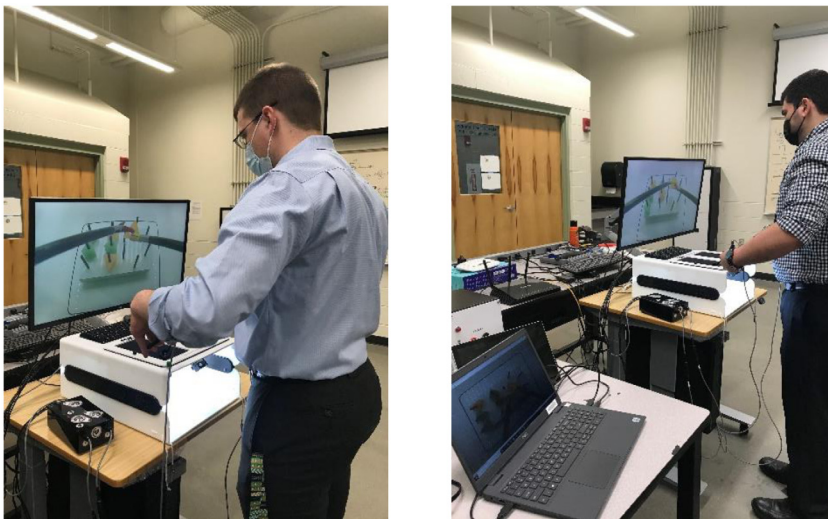


Figure 2. An overview of the proposed algorithm.

Table II. *The number of appearances of each instrument in the dataset.*

Instrument	Available size	Instrument	Available size
OnPeg	8000	Carry_l	4000
OutPeg	1500	Transfer	3000
Grasper_R	5000	Pick_R	4000
Grasper_L	5000	Pick_L	4000
Carry_R	4000	OutField	2500

**Figure 3.** *Carrying out peg transfer task by 9 WMed doctors and OB/GYN residents in the Intelligent Fuzzy Controllers Laboratory, WMU.*

according to the WMed doctors' expert knowledge on how to monitor sequentially each step of a surgeon's performance from three different cameras and assess whether the surgeon is performing the peg transfer task in the right order and display a notification on the monitor. The basic components of the peg transfer task include two surgical graspers, one pegboard with 12 pegs and six rubber triangle shape objects which are initially aligned on six pegs, on the same side of the pegboard.

2.3. Experimental peg transfer task for data collection

The experimental peg transfer task was carried out by nine WMed doctors and OB/GYN residents in the Intelligent Fuzzy Controllers Laboratory, WMU, as illustrated in Fig. 3. They possessed different levels of skill. To improve our first laparoscopic box-trainer dataset (IFCL.LBT100), 55 videos of peg transfer tasks were recorded using the top, side, and front cameras of our *n* IBTS in the Intelligent Fuzzy Controllers Laboratory, at WMU. More than 5000 images were extracted from those videos and were labeled manually by drawing a bounding box around the object by an expert. The number of each object label in the dataset is shown in Table II. Of all the images, 80% of them have been used for training and 20% for evaluation purposes.

2.4. Object detection and automatic tracking algorithm

The proposed methodology for the object detection is mainly divided into three parts: model training, model validation, and motion feature extraction and tracking point location. To monitor each step of a surgeons' performance sequentially, assess, and notify them immediately during the peg transfer task, a

robust object detection algorithm is proposed which is capable of recognizing and localizing each object correctly during the process and is summarized as follows:

- **Instrument and object types recognition:** the ability to distinguish types of tools and other objects when performing multi-object tracking and localization [37, 62].
- **Instrument tips and object localization:** giving the instruments' and objects' position as precisely as possible through the image acquired by the three cameras is a prerequisite for accurate tracking [63].
- **Real-time application:** Because MIS is a real-time process, peg transfer task should be done real time; hence, any unnecessary delay during operation is not acceptable during the procedures [64].

The SSD with ResNet50 V1 FPN feature extractor [65, 66] predicts the existence of an object in different scales based on several common features such as color, edge, centroid, and texture. Object tracking is a process of object feature detection, in which unique features of an object are detected. In this work, to detect and track the object based on the ResNet50 V1 FPN feature extractor, the related bounding box corresponding to a ground truth box should be found. All boxes are on different scales, and the best Jaccard overlap [67] should be found over a threshold >0.5 to simplify the learning problem.

2.4.1. Model training

In this study, the SSD ResNet50 V1 FPN Architecture, a feature extractor, has been trained on our custom dataset. The learning network of the SSD model is based on the backbone of a standard CNN architecture, designed for the semantic object classifier. A momentum optimizer with a learning rate of 0.04, which is not a very small and neither is a large one, was used for the region proposal and classification network. Obviously, a very small learning rate reduces the learning process but converges smoothly, and a large learning rate speeds up the learning but may not converge [14].

The input to the network is an image of the arbitrary size of 1280×720 pixels, and the frame rate is of 30 frames per second (FPS). The final output is a bounding box for each detected object as a class label (OnPeg, OutPeg, OutField, Grasper_R, Grasper_L, Carry_R, Carry_L, Transfer, Pick_R, and Pick_L) along with its confidence score. To carry out our experiments, the IBTS software was developed using the Python programming language in the Tensorflow deep learning platform [68]. The feasibility of this work has been evaluated using the TensorFlow API under Windows along with the NVIDIA GeForce GTX 1660 GPU.

2.4.2. Model validation

After training the model using our customized IFCL.LBT100 dataset, model validation was carried out on the test set, as the input images of the trained model. A loss function was used to optimize the model, and it is defined as the difference between the predicted value of the trained model and the true value of the dataset. The loss function used in this deep neural network is cross-entropy. It is defined as [14]:

$$\text{cross-entropy} = - \sum_{i=1}^n \sum_{j=1}^m y_{ij} \log(p_{ij}) \quad (1)$$

where y_{ij} denotes the true value, for example, 1 if sample i belongs to any of those 10 j classes, and 0 if it does not belong to any of them, and p_{ij} denotes the probability, predicted by the trained model that sample i belongs to any of those 10 j classes. It is illustrated in Fig. 4 what the network predicted for frames versus the allocated label of the image at the end of each epoch during the training process. The train-validation total loss at the end of 25,000 epochs for the SSD ResNet50 v1 FPN was about 0.05. Many object detection algorithms use the mean average precision (mAP) metric to analyze and evaluate the performance of the object detection in trained models. The mean of average precision (AP) values, which can be based on the intersection over union (IoU), are calculated over recall values from 0 to 1.

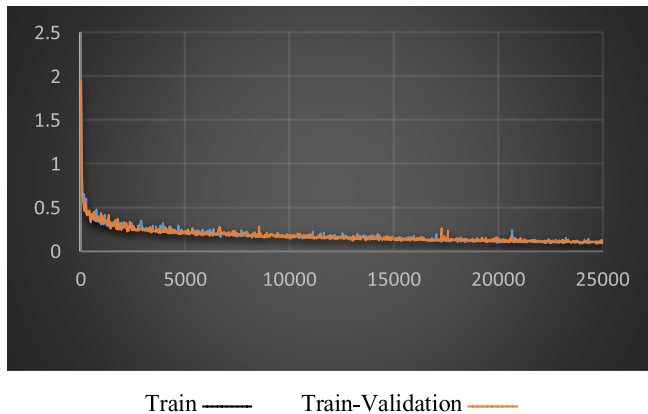


Figure 4. Train and train-validation total loss for SSD ResNet50 v1 FPN.

IoU designates the overlap of the predicted bounding box coordinates to the ground truth box. Higher IoU means the predicted bounding box coordinates closely resemble the ground truth box coordinates [69]. The IoU in this study is 0.75; therefore, based on ref. [70], the performance of this learning system is high enough. Also, the mAP has been used to measure the accuracy of the object detection model, which was reached at 0.741.

2.4.3. Motion feature extraction and tracking point location

The trained model runs inferences that return a dictionary with a list of bounding boxes' coordinates [xmax, xmin, ymax, and ymin] with prediction scores. To implement the tracking procedure, the tracking point must be located frame by frame in our laparoscopic videos. The output of the instrument/object detection models comprises the predicted instruments for every frame along with the x and y coordinates of their associated bounding boxes.

2.4.4. Sequential surgical skills assessment for the peg transfer task

The sequential surgical skill assessment algorithm for the peg transfer task has been developed based on several sequential If-Then conditional statements to monitor each step of a surgeon's performance sequentially from three different cameras, assess if the surgeon is performing the peg transfer task in the correct order, and display a notification instantly on the monitor. To perform a peg transfer task correctly, several constraints should be considered within 300 s. In this task, the time execution starts when the trainee touches the first object and ends upon the release of the last object. The flowchart of the hierarchical assessment algorithm is shown in Figs. 5 and 6, respectively. One of the main concepts of this algorithm is the introduction of several flags, such as carry right flag, related to an object carrying by the right grasper, carry left flag, related to an object carrying by the left grasper, carry flag, related to an object carrying, air flag, related to any grasper, carrying an object, and OutField flag, related to an object, which is dropped outside of the field – that is, determining the status of an object in the field or out of the field. Also, there are two counters, OnPeg counter, OutPeg counter that count the objects on the peg and the dropped ones inside the field, respectively. A Process List has been created to save any right and left movement, as well as transferring an object from the right grasper to the left grasper (vice versa) to provide a piece of precise information for making a proper decision on whether the transfer action has been done or not. Based on the flowcharts in Figs. 5 and 6, the following algorithm was implemented:

- a. Time starts
- b. OnPeg counter starts counting when an object is detected OnPeg.

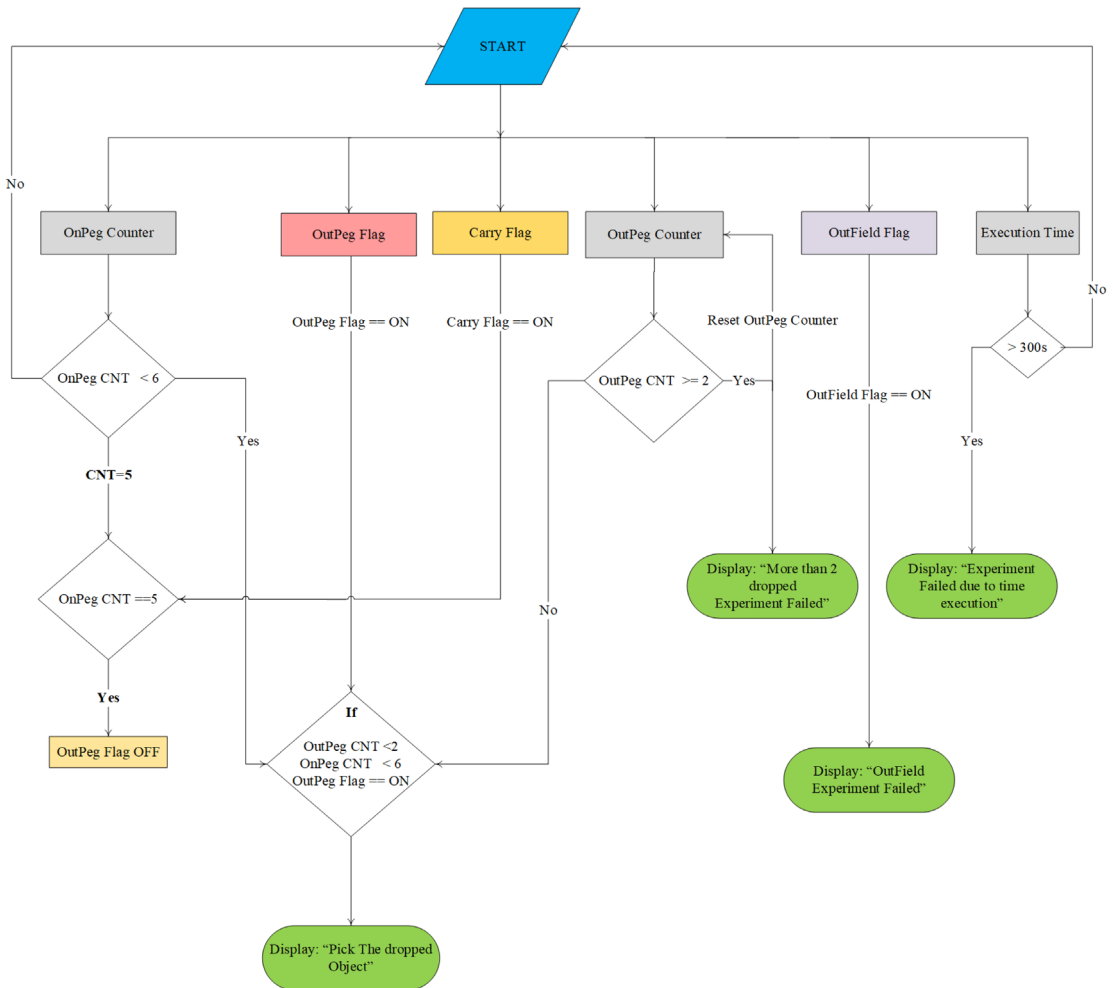


Figure 5. Sequential surgical skills assessment algorithm – results and display.

- c. If carry right flag is detected, “R” is added to the Process List and the carry flag is turned on.
- d. If carry left is detected, then “L” is added to the Process List, and the carry flag is turned on.
- e. If transferring is detected and the carry flag is one, then “T” is added to the Process List, and the carry flag is turned off.
- f. If R, T, L are in the Process List and the OnPeg counter is 6, then display “Transfer Done,” otherwise display “Transfer Failed”
- g. If any object drops within the field, the OutPeg counter starts counting.
- h. If the OutPeg counter counts 1 dropped object, display “Pick the dropped object.”
- i. If the OutPeg counter counts 2 dropped objects, display “More than 2 dropped objects: The experiment failed.”
- j. If any object is dropped outside of the field, the OutField flag turns on, display “OutField: The experiment failed.”
- k. If the task execution time exceeds 300 s, display “Time is more than 300 s, Failed Experiment.”
- l. Return to step (a)

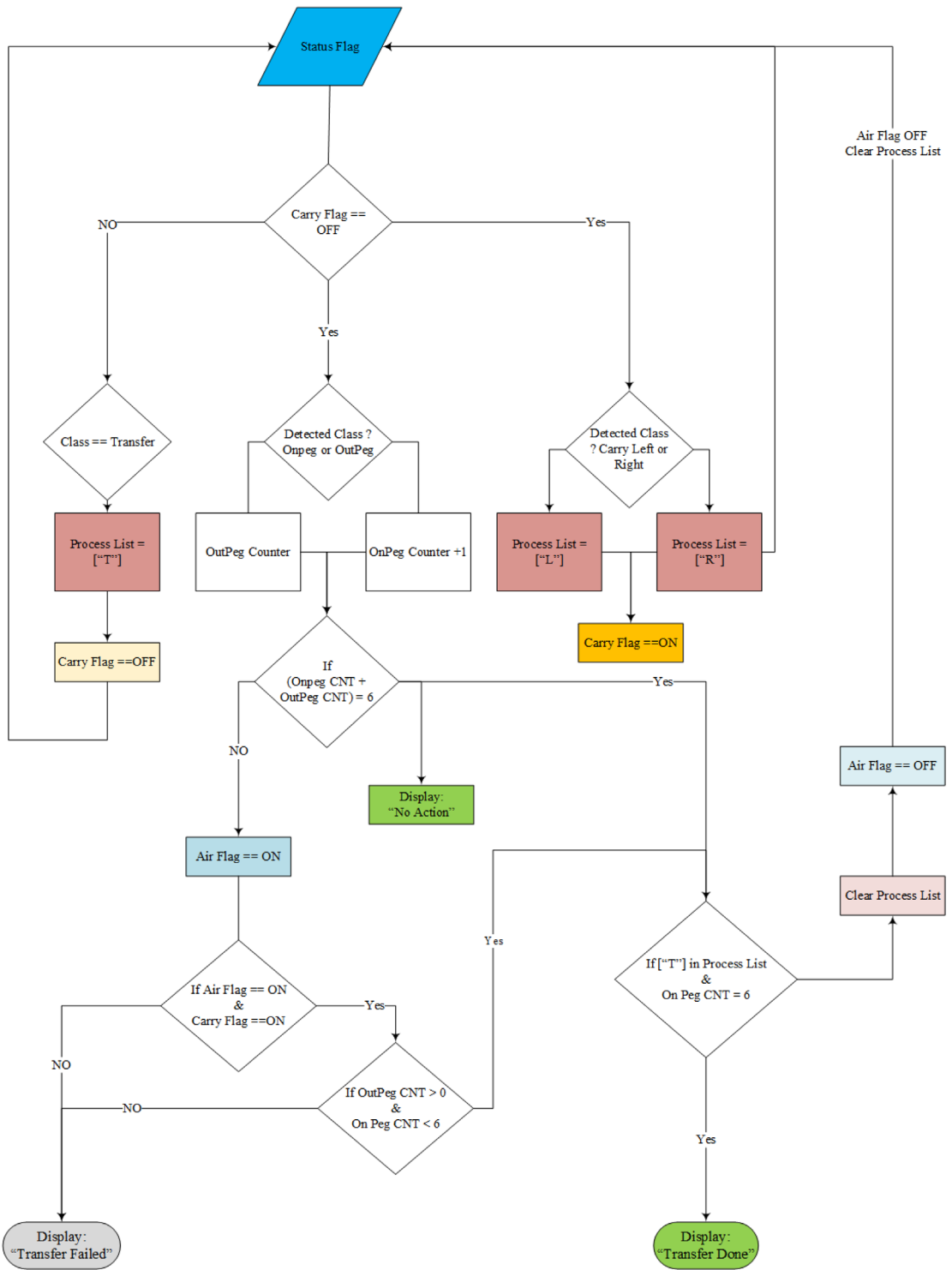


Figure 6. Sequential surgical skills assessment algorithm – internal analysis.

3. Experimental results

3.1. Results

The novelty of this work is about the autonomous, sequential assessment of the performance of a resident or a surgeon in the peg transfer task which also opens a new door to surgeons' performance assessment in robotic surgery. The proposed algorithm has the ability of detecting each peg transfer step separately and notifying the user on accomplished or failed outcome. To evaluate the accuracy of the trained model in the object detection section, evaluation metrics such as loss, mAP, and IoU were calculated. mAP is the value, from 0 to 1, of the area under the Precision–Recall curve. The value of AP varies from 0.52 to 0.79 and depends on the IoU threshold; therefore, an accurate learning system has a higher value of the IoU. In this study, the IoU threshold was set to 0.5 which is commonly used in other object detection studies [70]. The mAP was calculated by taking the average of AP over all the 10 classes. The results of the object detection section of this study demonstrated that the AP of each prediction varied from 0.52 to 0.79 with mAP of 0.741, and IoU was 0.75 with the train-validation total loss of 0.05.

In this section, the implementation details of the sequential surgical skills assessment system are illustrated in Fig. 7. In Fig. 7, a complete detection and autonomous sequential assessment procedure from the side camera is shown. In Figs. 8, 9, and 10, the assessment is shown only partially for the moment of acceptance or failure. As shown in Fig. 7, SSD ResNet50 v1 FPN could detect all the available objects in the field as well as the state of each object during the peg transfer task. In Fig. 7, all of the triangle objects had been placed in the left side of the pegboard, and the resident should have transferred each of these objects to the right side of the pegboard. The resident has already transferred one of the objects to the right side of the pegboard as depicted in Fig. 7(a). Meanwhile, the execution time is recorded and shown during the experiment. In Fig. 7(a), when the Carry_L class, standing for an object which is being carried on the left side by using the left grasper, is detected then the carry flag is turned on, and “L” will be added to the Process List. As soon as the OnPeg counter counts six objects on the peg, the notification of “Transfer Done” is displayed. In Fig. 7(b), the trainee is picking and carrying an object with the right grasper. When the Carry_R class is detected, “R” is added to the Process List. In Fig. 7(c), the Transfer class is detected, and “T” is added to the Process List. In Fig. 7(d), after detecting Carry_L class and adding “L” to the Process List, the OnPeg counter counts all the on-peg objects and the notification of “Transfer Done” is displayed.

When transferring an object from the left grasper to the right grasper is not detected, the “Transfer Failed” notification will be displayed. In this experiment, the resident transferred all objects in 50 s from the left side to the right side, and it took about 40 s to repeat the task in the opposite direction, as shown in Fig. 7(e)–(h). The performance of the trainee is shown from the front camera in Fig. 8. The resident had to pick up the object and place it on the peg and repeat the transfer and the placement process. As shown in Fig. 9, after 61 s, the object was dropped outside of the field. In this condition, the OutField class is detected and the OutField flag is turned on, and, because the experiment has failed, the notification of “OutField: The experiment is failed” is displayed on the screen. In Fig. 10, the assessment is shown from the top camera. As depicted in Fig. 10(b), when the object drops inside the field for the first time, the notification of “Pick up the dropped object” is shown on the screen. This assessment is based on the OutPeg class detection and three more parameters: if the OutPeg flag is on, and if the OutPeg Counter is less than 2, and if the OnPeg Counter is less than 6. If all these evaluation statements are met, then that notification will appear on the screen. In this particular case, the object was dropped inside the field after 44 s and the notification of “Pick up the dropped object” was shown.

3.2. Discussion

This method was used to assess the trainees' performance, after the operation was done, on the recorded videos. After implementing the algorithm on the FLS box trainer, the recorded videos are fed to the algorithm (implemented on our IBTS) to be evaluated. In the next phase of this study, the methodology will be examined under real-time conditions, while the trainee carrying out the peg transfer task and the

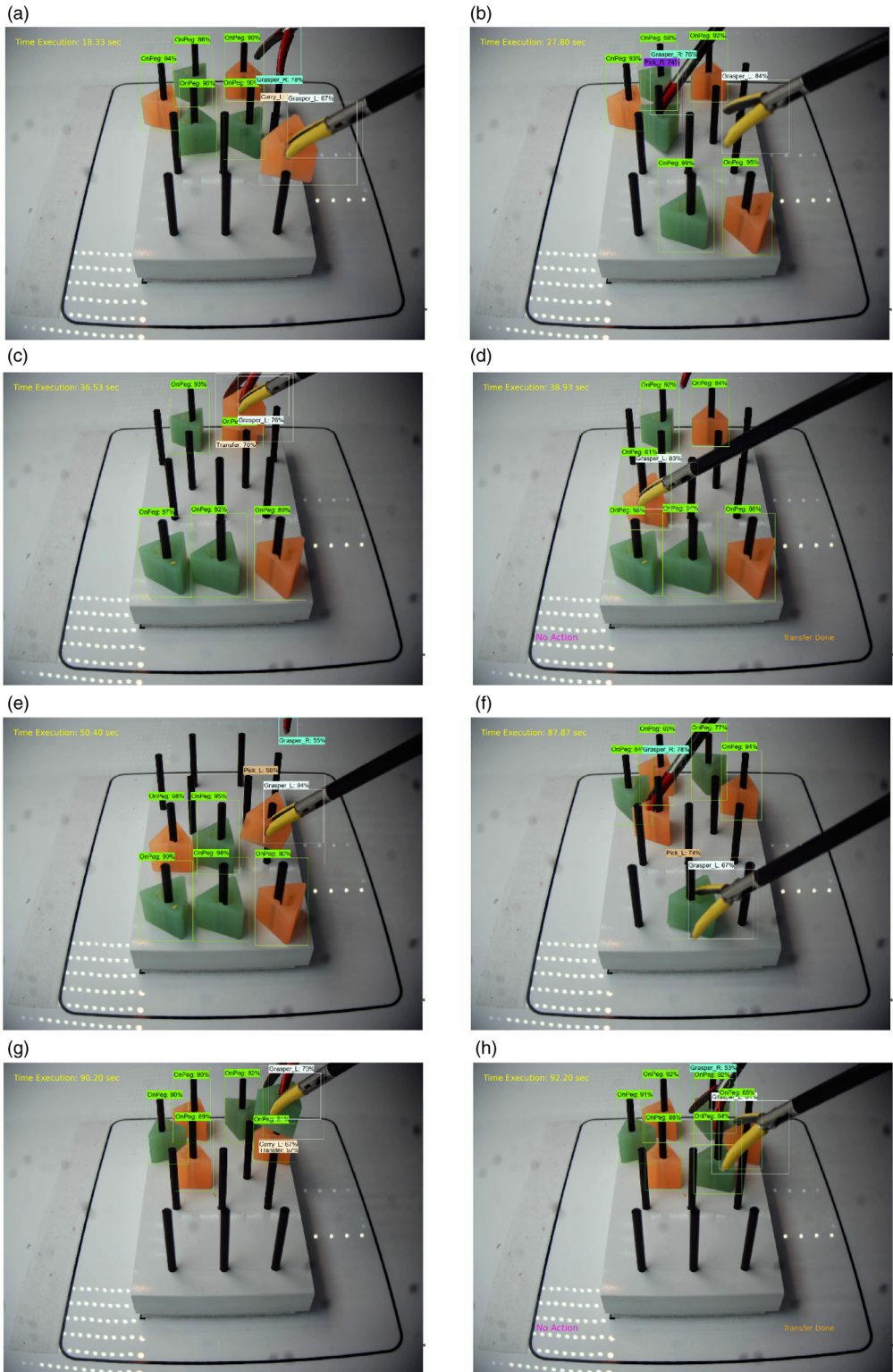


Figure 7. Peg transfer task assessment in a real operation using the sequential surgical skills assessment – side camera.

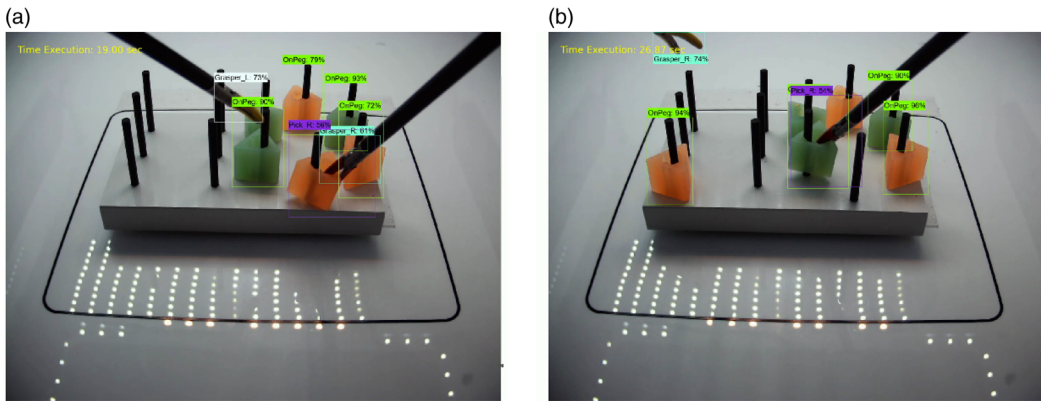


Figure 8. Peg transfer task assessment in a real operation using the sequential surgical skills assessment – front camera.

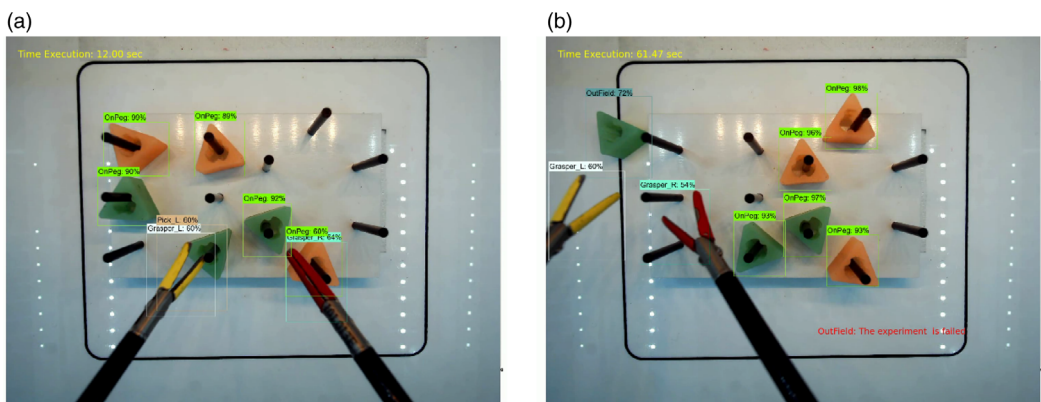


Figure 9. A failed Peg transfer task assessment – top Camera.

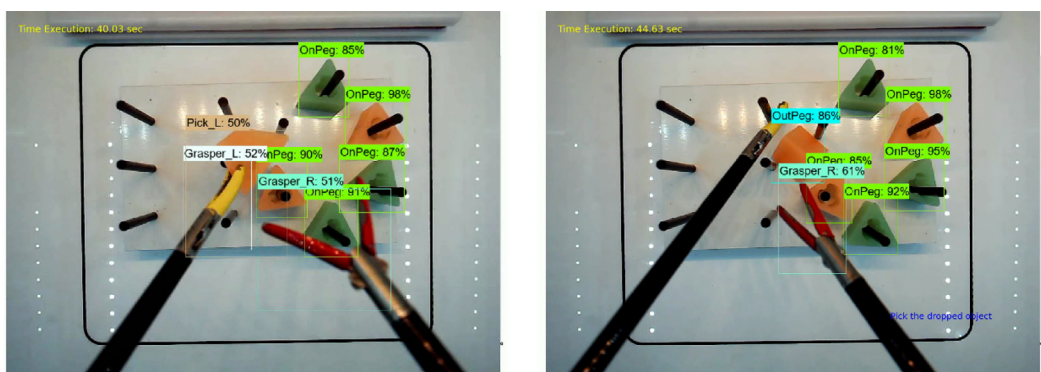


Figure 10. Repeating the peg transfer task after one time dropping of the object – top camera.

experiment assessment video is being recorded simultaneously. However, to accomplish all that is still a challenge. As mentioned before, currently we prefer to use the recorded assessment videos after the real-time experiment. That is due to the delay that the whole system exhibits, which is about 10 s in detection, because of the relatively small available memory and not having a better GPU. We plan to replace the PC in the IBTS with a more powerful computer, extend its memory to 32 GB, and install the NVIDIA

GeForce RTX 3080 10 GB GDDR6X 320-bit PCI-e Graphics Card in the PC to reduce this delay. We hope that with these improvements near real-time performance assessment can be achieved and have already developed a two-level fuzzy logic assessment system for faster real-time video processing using multi-threading on our IBTS system.

4. Conclusions

In laparoscopic surgery training, the peg transfer task is a hands-on exam in the FLS program which should be carried out by residents, and the results are assessed by medical personnel. In this study, an autonomous sequential algorithm is proposed to assess the resident's performance during the peg transfer task. A multi-object detection method which is based on the SSD ResNet50 v1 FPN architecture is proposed to improve the performance of the laparoscopic box-trainer-based skill assessment system by using top, side, and front mounted cameras. A sequential assessment algorithm was developed to detect each step of the peg transfer task separately and notify the resident of the accomplished, or failed performance. In addition, this method can correctly obtain the peg transfer execution time, the moving, carrying, and dropped states of each object from the top, side, and front cameras. During the assessment, the algorithm can detect the transfer of each object based on several defined counters and flags, and in case of a mistake, for example, failing to make the transfer between the two graspers, the trainee will be made aware of it. Also, if objects were dropped more than two times or if one object was dropped outside of the field, the notification of the failed performance would be given. The advantage of this autonomous, sequential skill assessment system is that even if a trainee or the expert that is evaluating the trainee failed to notice any mistakes, the system will record and report any bad performance. Based on the experimental results, the proposed surgical skill assessment system can identify each object at a high score of fidelity, and the train-validation total loss for the SSD ResNet50 v1 FPN was about 0.05. Also, mAP and IoU of this detection system are 0.741, and 0.75, respectively. As a conclusion, this proposed sequential assessment algorithm has the potential to provide fast, hierarchical, and autonomous feedback to the trainee without the presence of an expert medical personal for the FLS peg transfer task.

In future work, we plan to develop a multithreaded algorithm which will have the ability to assess the surgeons' performance based on the grasper tip movements in a virtual 3D space. This multi-threading algorithm will be implemented on our IBTS system. This project will include the upgrade of the whole system to a high performance one to achieve real-time assessment. Also, we would like to prepare a dataset using the residents' recorded performance in the peg transfer task to apply a machine learning algorithm which will predict if a resident is ready for a real laparoscopic surgery procedure, or not. For supporting research in the laparoscopic surgery skill assessment area, a Dataset has been created and made available, free of charge, at: <https://drive.google.com/drive/folders/1F97CvN3GnLj-rqg1tk2rHu8x0J740DpC>. As our research progresses, more files will be added to this Dataset. The code and the dataset will also be shared in the first authors' GitHub link (after the publication of the paper at: <https://github.com/Atena-Rashidi>).

Acknowledgments. This work was supported by Homer Stryker M.D. School of Medicine, WMU (Contract #:29-7023660), and the Office of Vice President for Research (OVPR), WMU (Project #: 161, 2018-19). We gratefully thank Mr. Hossein Rahmatpour (University of Tehran) for sharing his experience in this type of work, hosseinrahmatpour@ut.ac.ir, Mohsen Mohaidat, mohsen.mohaidat@wmich.edu for preparing the dataset, and Dr. Mohammed Al-Gailani for his contributions to the IBTS with three cameras. We also appreciate the doctors and residents who took part in this study: Dr. Saad Shebrain, Dr. Caitlyn Cookenmaster, Dr. Cole Kircher, Dr. Derek Tessman, Dr. Graham McLaren, Dr. Joslyn Jose, Dr. Kyra Folkert, Dr. Asma Daoudi, Dr. Angie Tsuei from Homer Stryker M.D. School of Medicine, WMU.

Author contributions. *Fatemeh Rashidi Fathabadi*: literature review, image processing and AI methods, fuzzy assessment system, database development.

Dr. Janos Grantner: research objectives and fuzzy assessment system.

Dr. Saad Shebrain: laparoscopic surgery skill assessment and arrangements for tests.

Dr. Ikhlas Abdel-Qader: image processing and AI methods.

Financial support. This work was supported by Homer Stryker M.D. School of Medicine, WMU (Contract #:29-7023660), and the Office of Vice President for Research (OVPR), WMU (Project #: 161, 2018-19).

Conflicts of interest. None.

Ethical approval. Yes.

References

- [1] J. Peng, Q. Chen, L. Kang, H. Jie and Y. Han, "Autonomous recognition of multiple surgical instruments tips based on arrow OBB-YOLO network," *IEEE Trans. Instrum. Meas.* **71**, 1–13 (2022).
- [2] K. N. Alkhamaiseh, J. L. Grantner, S. Shebrain and I. Abdel-Oader, "Towards Automated Performance Assessment for Laparoscopic Box Trainer using Cross-Stage Partial Network," *In: 2021 Digital Image Computing: Techniques and Applications (DICTA)* (2021) pp. 1–7.
- [3] S. Ohtake, K. Makiyama, D. Yamashita, T. Tatenuma and M. Yao, "Training on a virtual reality laparoscopic simulator improves performance of live laparoscopic surgery," *Asian J. Endosc. Surg.* **15**(2), 313–319 (2022).
- [4] Y. Hiyoshi, Miyamoto Y., Akiyama T., Daitoku N., Sakamoto Y., Tokunaga R., Eto K., Nagai Y., Iwatsuki M., Iwagami S., Baba Y., Yoshida N. and Baba H., "Time trial of dry box laparoscopic surgical training improves laparoscopic surgical skills and surgical outcomes," *Asian J. Endosc. Surg.* **14**(3), 373–378 (2021).
- [5] J. Davids, S.-G. Makariou, H. Ashrafiyan, A. Darzi, H. J. Marcus and S. Giannarou, "Automated vision-based microsurgical skill analysis in neurosurgery using deep learning: Development and preclinical validation," *World Neurosurg.* **149**, e669–e686 (2021).
- [6] A. Maciel, Y. Liu, W. Ahn, T. P. Singh, W. Dunnican and S. De, "Development of the VBLaST™: A virtual basic laparoscopic skill trainer," *Int J. Med. Robot. Comput. Assist. Surg.* **4**(2), 131–138 (2008).
- [7] O. J. Y. Kirsten, T. W. K. Andy, D. F. C. Wei and L. S. K. Jason, "Simulation in surgical education: A review of a multi-tiered gynaecological laparoscopy workshop," *South-East Asian J. Med. Educ.* **15**(1), 34 (2021).
- [8] S. Matsumoto, H. Kawahira, K. Oiwa, Y. Maeda, A. Nozawa, A. K. Lefor, Y. Hosoya and N. Sata, "Laparoscopic surgical skill evaluation with motion capture and eyeglass gaze cameras: A pilot study," *Asian J. Endosc. Surg.* **15**(3), 619–628 (2022).
- [9] N. Ahmidi, P. Poddar, J. D. Jones, S. S. Vedula, L. Ishii, G. D. Hager and M. Ishii, "Automated objective surgical skill assessment in the operating room from unstructured tool motion in septoplasty," *Int J. Comput. Assist. Radiol. Surg.* **10**(6), 981–991 (2015).
- [10] J. B. Pagador, F. M. Sánchez-Margallo, L. F. Sánchez-Peralta, J. A. Sánchez-Margallo, J. L. Moyano-Cuevas, S. Enciso-Sanz, J. Usón-Gargallo and J. Moreno, "Decomposition and analysis of laparoscopic suturing task using tool-motion analysis (TMA): Improving the objective assessment," *Int. J. Comput. Assist. Radiol. Surg.* **7**(2), 305–313 (2012).
- [11] S. O'Sullivan, N. Nevejans, C. Allen, A. Blyth, S. Leonard, U. Pagallo, K. Holzinger, A. Holzinger, M. I. Sajid and H. Ashrafiyan, "Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery," *Int J. Med. Robot. Comput. Assist. Surg.* **15**(1), e1968 (2019).
- [12] S. Suzuki, Suzuki N., Hattori A., Hayashibe M., Konishi K., Kakeji Y. and Hashizume M., "Tele-surgery simulation with a patient organ model for robotic surgery training," *Int. J. Med. Robot. Comput. Assist. Surg.* **1**(4), 80–88 (2005).
- [13] K. Song and P. Hsieh, "Shared-control design for robot-assisted surgery using a skill assessment approach," *Asian J. Control* **24**(3), 1042–1058 (2022).
- [14] E. De Ravin, E. A. Sell, J. G. Newman and K. Rajasekaran, "Medical malpractice in robotic surgery: A Westlaw database analysis," *J. Robot. Surg.*, 1–6 (2022).
- [15] M. A. Zeinab and J. Kaouk, "Novel Technology in Robotic Surgery," *In: Atlas of Robotic, Conventional, and Single-Port Laparoscopy* (Springer, Cham, 2022) pp. 247–257.
- [16] L.-W. Sun, F. Van Meer, J. Schmid, Y. Bailly, A. A. Thakre and C. K. Yeung, "Advanced da Vinci surgical system simulator for surgeon training and operation planning," *Int. J. Med. Robot. Comput. Assist. Surg.* **3**(3), 245–251 (2007).
- [17] F. Qin, S. Lin, Y. Li, R. A. Bly, K. S. Moe and B. Hannaford, "Towards better surgical instrument segmentation in endoscopic vision: Multi-angle feature aggregation and contour supervision," *IEEE Robot. Autom. Lett.* **5**(4), 6639–6646 (2020).
- [18] Z. Zhao, Jin Y., Lu B., Ng C.-F., Dou Q., Liu Y.-H. and Heng P.-A., "One to Many: Adaptive Instrument Segmentation via Meta Learning and Dynamic Online Adaptation in Robotic Surgical Video," *In: 2021 IEEE International Conference on Robotics and Automation (ICRA)* (2021) pp. 13553–13559.
- [19] H. Nakawala, R. Bianchi, L. E. Pescatori, O. De Cobelli, G. Ferrigno and E. De Momi, "'Deep-Onto' network for surgical workflow and context recognition," *Int. J. Comput. Assist. Radiol. Surg.* **14**(4), 685–696 (2019).
- [20] B. van Amsterdam, M. J. Clarkson and D. Stoyanov, "Gesture recognition in robotic surgery: A review," *IEEE Trans. Biomed. Eng.* **68**(6), 2021–2035 (2021).
- [21] J. Zhang, Y. Nie, Y. Lyu, X. Yang, J. Chang and J. J. Zhang, "SD-Net: Joint surgical gesture recognition and skill assessment," *Int. J. Comput. Assist. Radiol. Surg.* **16**(10), 1675–1682 (2021).
- [22] Y. Long, Li Z., Yee C. H., Ng C. F., Taylor R. H., Unberath M. and Dou Q., "E-DSSR: Efficient Dynamic Surgical Scene Reconstruction with Transformer-Based Stereoscopic Depth Perception," *In: International Conference on Medical Image Computing and Computer-Assisted Intervention* (2021) pp. 415–425.

- [23] Y. Long, J. Cao, A. Deguet, R. H. Taylor and Q. Dou, "Integrating Artificial Intelligence and Augmented Reality in Robotic Surgery: An Initial dVRK Study Using a Surgical Education Scenario," *arXiv Preprint. arXiv:2201.00383* (2022).
- [24] J. Han, J. Davids, H. Ashrafian, A. Darzi, D. Elson and M. Sodergren, "A systematic review of robotic surgery: From supervised paradigms to fully autonomous robotic approaches," *Int. J. Med. Robot. Comput. Assist. Surg.* **18**(2), e2358 (2021).
- [25] G. T. Gonzalez, U. Kaur, M. Rahman, V. Venkatesh, N. Sanchez, G. Hager, Y. Xue, R. Voyles and J. Wachs, "From the dexterous surgical skill to the battlefield—a robotics exploratory study," *Mil. Med.* **186**(Supplement_1), 288–294 (2021).
- [26] J. Rosen and J. Ma, "Autonomous operation in surgical robotics," *Mech. Eng.* **137**(9), S15–S18 (2015).
- [27] B. Hannaford, J. Rosen, D. W. Friedman, H. King, P. Roan, Lei Cheng, D. Glozman, Ji Ma, S. N. Kosari and L. White, "Raven-II: An open platform for surgical robotics research," *IEEE Trans. Biomed. Eng.* **60**(4), 954–959 (2012).
- [28] M. Hwang, Seita D., Thananjeyan B., Ichnowski J., Paradis S., Fer D., Low T. and Goldberg K. Y., "Applying Depth-Sensing to Automated Surgical Manipulation with a da Vinci Robot," *In: 2020 International Symposium on Medical Robotics (ISMR)* (2020) pp. 22–29.
- [29] M. Hwang, Thananjeyan B., Seita D., Ichnowski J., Paradis S., Fer D., Low T. and Goldberg K. Y., "Superhuman Surgical Peg Transfer Using Depth-Sensing and Deep Recurrent Neural Networks," *arXiv Preprint. arXiv:2012.12844* (2020).
- [30] Z. Qin, Y. Tai, C. Xia, J. Peng, X. Huang, Z. Chen, Q. Li and J. Shi, "Towards virtual VATS, face, and construct evaluation for peg transfer training of box, VR, AR, and MR trainer," *J. Healthc. Eng.* **2019**, 1–10 (2019).
- [31] J. D. Brown, C. E. O'Brien, S. C. Leung, K. R. Dumon, D. I. Lee and K. J. Kuchenbecker, "Using contact forces and robot arm accelerations to automatically rate surgeon skill at peg transfer," *IEEE Trans. Biomed. Eng.* **64**(9), 2263–2275 (2016).
- [32] C. Aschwanden, L. Burgess and K. Montgomery, "Performance Compared to Experience Level in a Virtual Reality Surgical Skills Trainer," *In: International Conference on Foundations of Augmented Cognition* (2007) pp. 394–399.
- [33] M. M. Coad, Okamura A. M., Wren S., Mintz Y., Lendvay T. S., Jarc A. M. and Nisky I., "Training in Divergent and Convergent Force Fields During 6-DOF Teleoperation with a Robot-Assisted Surgical System," *In: 2017 IEEE World Haptics Conference (WHC)* (2017) pp. 195–200.
- [34] J. Chen, Zhang D., Munawar A., Zhu R., Lo B., Fischer G. S. and Yang G.-Z., "Supervised Semi-Autonomous Control for Surgical Robot Based on Banoian Optimization," *In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2020) pp. 2943–2949.
- [35] A. R. Hutchins, R. J. Manson, S. Zani and B. P. Mann, "Sample Entropy of Speed Power Spectrum as a Measure of Laparoscopic Surgical Instrument Trajectory Smoothness," *In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2018) pp. 5410–5413.
- [36] Z. Zhao, S. Voros, Y. Weng, F. Chang and R. Li, "Tracking-by-detection of surgical instruments in minimally invasive surgery via the convolutional neural network deep learning-based method," *Comput. Assist. Surg.* **22**(sup1), 26–35 (2017).
- [37] Z. Zhao, Z. Chen, S. Voros and X. Cheng, "Real-time tracking of surgical instruments based on spatio-temporal context and deep learning," *Comput. Assist. Surg.* **24**(sup1), 20–29 (2019).
- [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* **19**(2), 7–27 (2022).
- [39] 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.* **15**(8), 1335–1345 (2020).
- [40] K. Kunkler, "The role of medical simulation: An overview," *Int. J. Med. Robot. Comput. Assist. Surg.* **2**(3), 203–210 (2006).
- [41] 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.
- [42] Y. Gao, U. Kruger, X. Intes, S. Schwaitzberg and S. De, "A machine learning approach to predict surgical learning curves," *Surgery* **167**(2), 321–327 (2020).
- [43] R. J. Kuo, H.-J. Chen and Y.-H. Kuo, "The development of an eye movement-based deep learning system for laparoscopic surgical skills assessment," *Sci. Rep.* **12**(1), 1–12 (2022).
- [44] M. Hong, K. Meisner, S. Lee, A. M. Schreiber and J. W. Rozenblit, "A Fuzzy Reasoning System for Computer-Guided Laparoscopy Training," *In: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (2020) pp. 1712–1717.
- [45] G. Islam, K. Kahol, B. Li, M. Smith and V. L. Patel, "Affordable, web-based surgical skill training and evaluation tool," *J. Biomed. Inform.* **59**, 102–114 (2016).
- [46] K. S. Peng, M. Hong, J. Rozenblit and A. J. Hamilton, "Single Shot State Detection in Simulation-Based Laparoscopy Training," *In: 2019 Spring Simulation Conference (SpringSim)* (2019) pp. 1–12.
- [47] H. Jiang, Xu S., State A., Feng F., Fuchs H., Hong M. and Rozenblit J., "Enhancing a Laparoscopy Training System with Augmented Reality Visualization," *In: 2019 Spring Simulation Conference (SpringSim)* (2019) pp. 1–12.
- [48] K. Meisner, M. Hong and J. W. Rozenblit, "An Object State Estimation for the Peg Transfer Task in Computer-Guided Surgical Training," *In: 2020 Spring Simulation Conference (SpringSim)* (2020) pp. 1–12.
- [49] C. Harris and M. Stephens, "A combined corner and edge detector," *Alvey Vis. Conf.* **15**(50), 10–5244 (1988).
- [50] K. S. Peng, M. Hong and J. Rozenblit, "Image-Based Object State Modeling of a Transfer Task in Simulated Surgical Training," *In: Proceedings of the Symposium on Modeling and Simulation in Medicine* (2017) pp. 1–12.
- [51] F. Pérez-Escamirosa, I. Oropesa, P. Sánchez-González, J. Tapia-Jurado, J. Ruiz-Lizarraga and A. Minor-Martínez, "Orthogonal cameras system for tracking of laparoscopic instruments in training environments," *Cir. Cir.* **86**(6), 548–555 (2019).

- [52] I. Oropesa, P. Sánchez-González, M. K. Chmarra, P. Lamata, Álvaro Fernández, J. A. Sánchez-Margallo, F. W. Jansen, J. Dankelman, F. M. Sánchez-Margallo and E. J. Gómez, “EVA: Laparoscopic instrument tracking based on endoscopic video analysis for psychomotor skills assessment,” *Surg. Endosc.* **27**(3), 1029–1039 (2013).
- [53] B. F. Allen, F. Kasper, G. Nataneli, E. P. Dutton and P. Faloutsos, “Visual Tracking of Laparoscopic Instruments in Standard Training Environments,” *In: MMVR* (2011) pp. 11–17.
- [54] M. Chauhan, R. Sawhney, C. F. Da Silva, N. Aruparayil, J. Gnanaraj, S. Maiti, A. Mishra, A. Quyn, W. Bolton, J. Burke, D. Jayne and P. Valdastrì, “Evaluation and usability study of low-cost laparoscopic box trainer Lap-Pack: A 2-stage multicenter cohort study,” *IJS Glob. Health* **4**(5), e59 (2021).
- [55] S. Bao, J. Pan, X. Chang, D. Wu and C. Fang, “Virtual Surgical Instruments and Surgical Simulation,” *In: Biliary Tract Surgery* (Springer, Singapore, 2021) pp. 131–159.
- [56] K. Lam, J. Chen, Z. Wang, F. M. Iqbal, A. Darzi, B. Lo, S. Purkayastha and J. M. Kinross, “Machine learning for technical skill assessment in surgery: A systematic review,” *NPI Digit. Med.* **5**(1), 1–16 (2022).
- [57] H. Wang and X. Zhang, “Real-time vehicle detection and tracking using 3D LiDAR,” *Asian J. Control* **24**(3), 1459–1469 (2022).
- [58] H. A. Abbas, “A new adaptive deep neural network controller based on sparse auto-encoder for the antilock braking system systems subject to high constraints,” *Asian J. Control* **23**(5), 2145–2156 (2021).
- [59] J. Kim, “Autonomous rover guidance and localization by measuring the peak of a tall landmark,” *Asian J. Control* **24**(5), 2140–2152 (2022).
- [60] 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.
- [61] F. R. Fathabadi, J. L. Grantner, S. A. Shebrain and I. Abdel-Qader, “Multi-Class Detection of Laparoscopic Instruments for the Intelligent Box-Trainer System Using Faster R-CNN Architecture,” *In: 2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMII)* (2021) pp. 149–154.
- [62] L. Li, X. Li, B. Ouyang, S. Ding, S. Yang and Y. Qu, “Autonomous multiple instruments tracking for robot-assisted laparoscopic surgery with visual tracking space vector method,” *IEEE/ASME Trans. Mechatron.* **27**(2), 733–743 (2021).
- [63] 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)* (2018) pp. 624–628.
- [64] M. Robu, A. Kadkhodamohammadi, I. Luengo and D. Stoyanov, “Towards real-time multiple surgical tool tracking,” *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **9**(3), 279–285 (2021).
- [65] F. R. Fathabadi, J. L. Grantner, S. A. Shebrain and I. Abdel-Qader, “Fuzzy logic supervisor-A surgical skills assessment system using multi-class detection of laparoscopic box-trainer instruments,” *J. Intell. Fuzzy Syst.* **43**(4), 1–16 (2022).
- [66] F. R. Fathabadi, J. L. Grantner, S. A. Shebrain and I. Abdel-Qader, “Surgical Skill Training and Evaluation for a Peg Transfer Task in a Three Camera Based-Laparoscopic Box-Trainer System,” *In: 2021 International Conference on Computational Science and Computational Intelligence (CSCI)* (2021) pp. 1146–1151.
- [67] 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.
- [68] K. He, X. Zhang, S. Ren and J. Sun, “Deep Residual Learning for Image Recognition,” *In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016) pp. 770–778.
- [69] C. Wilms, A. M. Gerlach, R. Schmitz and S. Frintrop, “Segmenting Medical Instruments in Minimally Invasive Surgeries using AttentionMask,” *arXiv Preprint. arXiv:2203.11358* (2022).
- [70] T. Takahashi, K. Nozaki, T. Gonda, T. Mamenno and K. Ikebe, “Deep learning-based detection of dental prostheses and restorations,” *Sci. Rep.* **11**(1), 1–7 (2021).