Computerized model for objectively evaluating cutting performance using a laparoscopic box trainer simulator

Amir Handelman[1] · Shani Schnaider[2] · Adva Schwartz-Ossad[2] · Refael Barkan[1] · Ronnie Tepper[2,3]

Abstract

Background Laparoscopic box trainer simulator has recently become a tool for assessment of physicians' surgical and laparoscopic skills, and training using such a simulator has been incorporated into the curricula of surgery syllabus. With the increased use of box trainer simulators, there is a great need for obtaining reliable and objective evaluations of the trainees' performances. Here, we introduce an automated tool for assessing laparoscopic cutting performance by using image-processing algorithms.

Methods Twenty-seven interns specializing in the fields of gynecology, urology and general surgery participated in 4–6 training sessions, in which each trainee cut a circular patch positioned inside a low-cost laparoscopic box trainer simulator. The trainees' performances were analyzed using software that we developed. The analysis of the trainees' performances was based upon quantitative measurements of the following four parameters obtained in each training session: standard deviation, circle-cutout area, skewness, and number of peaks. We believe that high performance in terms of a *combination* of the four parameters provides a reliable measure of good laparoscopic skills, and therefore we developed the software so as to generate, for each session, a score of a trainee's laparoscopic circle-cutout performance that results from achievements related to the four parameters in combination.

Results On average, the total score of experienced interns was higher than the total score of inexperienced interns. Also, some improvement from session-to-session in the scores of novice trainees was detected.

Conclusions Our proposed scoring system, which is based on various image-processing algorithms, can evaluate cutting performances of trainees and classify residents by their experience. This allows each trainee to improve his/her performance by analyzing errors indicated by our software.

Keywords Laparoscopic box trainer simulator · Image-processing · Boundary extraction · Signature · Simulation · Surgery education

✉ Amir Handelman
  handelmana@hit.ac.il

✉ Ronnie Tepper
  tepper@clalit.org.il

1 Department of Electrical Engineering, Faculty of Engineering, Holon Institute of Technology, Holon, Israel

2 Simultech Center for Medical Simulation, Clalit Health Services, Meir Hospital Medical Center, Kfar Saba, Israel

3 Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

Today, there is an increasing desire by patients and hospital managers to make laparoscopic simulation trainings for residents and surgeons mandatory before performing actual operations [1], and indeed modern surgery syllabus includes use of simulators for improving laparoscopic surgical skills in teaching programs. A wide range of surgical procedures is now performed laparoscopically, including 90% of cholecystectomies in the United States [2], and 95% of pediatric appendectomies [3]. Most gynecologic adnexal procedures are performed laparoscopically, with the added advantage of causing fewer adhesions and decreasing fertility-related concerns [4, 5].

Complications arising from laparoscopic procedures typically include injuries during entry to the specific body cavity, injuries to adjacent organs and blood vessels, misuse of

electrocautery, and inaccurate cutting [6, 7]. Box-trainers and virtual reality-type simulations were developed [8–10] to overcome these issues.

The fundamentals of laparoscopic surgery (FLS) initiative was introduced in 2004 to provide a basic standard of training for those performing laparoscopic procedures [11–13]. For this purpose, different schemes of box trainer simulators (BTSs) were designed [14–22] and tested on junior and senior residents. The BTS is a relatively low-cost simulator that provides the experience of using actual laparoscopic devices. It usually includes five FLS training modules: peg transfer, pattern cutting, ligating loop, suturing with extracorporeal knot-tying procedures in the well-established McGill inanimate system for training and evaluation of laparoscopic skills (MISTELS), and suturing with intracorporeal knot-tying [13]. The task of sectioning a patch along a pre-designed circular template is one of the standard evaluations of the laparoscopic skills program [15].

Various studies showed that FLS simulator trainings have positive effects on residents' surgical and laparoscopic skills [15–17]. Consequently, such trainings are being incorporated into the curricula [11–15]. With the increased use of laparoscopic training, low-cost, easy-to-operate BTSs that generate reliable evaluations of the trainees are needed.

There are several ways to assess cutting performance when using laparoscopic simulators. Many grading methods rely on self-assessment questionnaires [13, 19, 20] or require the presence of a professional evaluator to provide constructive feedback [22]. By relying on human intervention, BTSs lack the tools for automatic, objective assessments.

BTSs that base evaluations on objective parameters, such as motion detection have also been proposed. Such parameters aim to assess surgical performance by measuring the surgeon's hand motion quality and smoothness with tracking systems [21, 23] or accelerometers [24]. While such BTS technology focuses on psychomotor skills, it does not generate objective reports on cutting or suturing accuracy. It should be noted that the studies which analyzed the motion quality and smoothness [23, 24] were focused merely on motion analysis parameters (MAP) for eye-hand coordination tasks [23] or on jerk index for micromanipulation tasks [24], and did not include any reference to the quality of the task performed, such as in terms of number and type of errors. Although evaluation of hand motion quality and smoothness provides valuable feedback during motion, it does not provide a total outcome of the procedure that can be viewed and analyzed for feedback after termination of a training session. Therefore, in order to get a full picture of the surgeon's performance it is also needed to quantify the outcomes of the trainees' different tasks. To meet this need, we developed the cut analysis laparoscopic simulator (CALS) software that uses a computerized algorithm to score the performance of dissecting a circle shape while using a low-cost BTS.

Our study introduces a novel objective tool for assessing laparoscopic cutting. It analyzes images of circular cuts (obtained by a simple web-camera) and generates their unique signature. In order to produce the image, the circle-cutout patch is inserted into a sealed chamber containing a camera. In contrast to the global operative assessment of laparoscopic skills (GOALS) tool [25], we developed autonomous computer software that generates a trainee's score based on a combination of objective parameters that represent the accuracy of the cut, including standard deviation, inner and outer excursions, skewness, and area. Using this algorithm, human assessment is not necessary, and the quality of the simulated laparoscopic surgery is automatically and objectively evaluated.

Materials and methods

Method

The study was conducted at the facilities of Simultech—The Unit for Medical Simulation of Clalit Health Services, Meir Hospital Medical Center, Kfar Saba, Israel. Medical residents specializing in gynecology, urology and general surgery participated in several series of training sessions under the guidance of an instructor. The training sessions allowed the residents to practice laparoscopic skills such as eye-hand coordination, cutting, intra-corporal knot-tying, and safe and efficient suturing. This work focused only on the exercise of cutting a circle shape using laparoscopic scissors and a Maryland grasper. The outcomes were objectively and quantitatively evaluated using the CALS software (IRB approval is not needed).

To assess trainees progress, performance parameters of relative area, standard deviation, number of excursions and skewness of the circle-cutout image (detailed in the section Objective parameters for assessing trainees' performance) were documented and the score given by the CALS software was calculated and determined by comparison to a reference circle-cutout. Furthermore, each performance score was compared to previous attempts of the same trainee; thus, a specific study-curve for each trainee was built.

Additional information on the trainees

It is assumed that during the learning process, parameters other than those evaluated in the training itself might affect performance. The experience in performing laparoscopic surgeries, field of internship, time to task completion and even sleeping time of the trainees before the training were

also recorded, but *not* considered when evaluating performance. These parameters appear in Table S1.

The simulator

The simulator used by Simultech is a BTS that was custom-made for Clalit Health Services. The simulator allows a trainee to practice laparoscopic skills and procedures using actual operating room (OR) instruments and OR optics, accessible through 9 ports.

Training

Interns specializing in the fields of gynecology, urology and general surgery, participated in three training series of increasing difficulty (basic, advanced and complicated). Each training series was composed of 4–6 sessions. This study focused on the first series, which addressed basic laparoscopic skills that include cutting a circle shape positioned inside the BTS.

Circle-cutout procedure

Figure 1 depicts the steps of cutting out a circle. First, a trainee dissects a circle shape (marked on a yellow patch with dimensions of $11.5 \times 12.5 \times 0.2$ cm) by using laparoscopic scissors and a Maryland grasper, as indicated in Fig. 1A, B. An instructor measures the completion time of the task, and supervises the trainee during the procedure. After the trainee finishes cutting out the circle, the instructor inserts the circle-cutout into the sealed chamber (Fig. 1C),
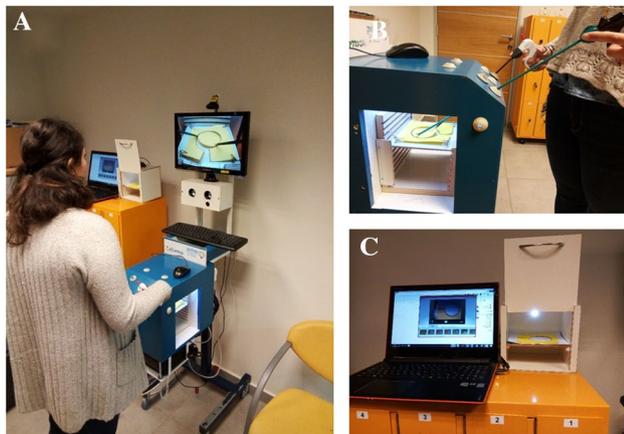

Fig. 1 Circle-cutout task performed according to the fundamentals of laparoscopic surgery (FLS) standards using a box trainer simulator (BTS). **A** System overview—performing a circle-cutout task and observing it on the monitor. **B** Close-up side-view of the circle pattern in the BTS. **C** Accuracy assessment sub-system. After completing the task, the circle-cutout is inserted into a sealed box with uniform lighting conditions, where it is photographed by a camera

and photographs it with a camera mounted at the top of the chamber. The circle-cutout photo is communicated to the computer using USB or Wi-Fi. The CALS software acquires the image and instantly produces a score ranging from 0 to 100. In this work we analyzed a total of 127 circle-cutout images. It is to be noted that the insertion of the circle-cutout patch into the chamber and measurement of the completion time are simple procedures that do not need any special training and may be performed either by the instructor or by the trainee. Assessment of the trainee's performance is provided by the software after the image is produced.

Analysis of circle-cutouts by CALS

In the first stage, the image with the circle-cutout is captured by the suspended camera. The second stage contains segmentation of the circle shape in the image. Next, the circle's boundaries are obtained by different morphological operations [26]. The final stage extracts the signature of the circle-cutout image. The signature is a representation of two-dimensional (2D) shapes as a one-dimensional (1D) function [26]. Once the circle-cutout center points are determined [27], the 1D representation function of the circle-cutout boundaries is obtained by measuring the Euclidean distance from the centroid of the circle [27] to every point on its boundary as a function of the angle (Appendix).

Although the signature of shapes is not invariant to scale and rotation [26], the cutting assignment is done on patches having constant circle sizes. Furthermore, since the camera is fixed at the top of a sealed chamber, the size of the image of every circle-cutout remains the same. As such, the circle-cutout scale is approximately constant and equals one. The rotation of the circle has little significance, because the shape is circular.

The signature of every trainee's circle-cutout is compared to the signature of a reference, optimal circle-cutout, which was obtained outside the BTS and simulates best performance. The parameters of the reference circle-cutout are expressed in the score formula (Eq. 1).

Objective parameters for assessing trainees' performance

The boundaries of the circle-cutout images are extracted by the CALS software. All parameters related to the boundaries are objective, and describe the accuracy of the cuts. Here, we describe four parameters that affect the final score of each trainee:

1. Standard deviation (SD or σ): The SD of the 1D boundaries function is a measure of the inner and outer deviations/errors from the perfect circle's contour (radius) and is measured for every circle-cutout image, in pixels

(Appendix). A good circle-cutout should have a small SD, that is, the SD should be as close as possible to the reference SD. The SD is an indication of the trainee's control of the laparoscopic scissors and grasper and of smoothness of movement during cutting.

2. Skewness (S): Skewness is a measure of the asymmetry of the radius. A good circle-cutout should have skewness close to zero (reference circle cut is 0.036). The skewness measures the "direction" of mistakes, i.e., if the trainee cuts are more toward the inside or the outside of the circle boundaries.

3. Number of peaks (NP): Peaks are local maxima (for excursions outside the circle) or local minima (for excursions towards the centroid of the circle) of the boundaries. The total number of peaks is the sum of the number of local maxima and local minima. A good circle-cutout should have zero peaks. The number of peaks can also indicate the amount of control with the laparoscopic scissors and grasper. The amplitude of the peaks is a measure of cutting inaccuracy, i.e., the deviation of the cut from the circle shape.

4. Area (A): The area (in pixels) of the circle-cutout is the integral below the boundaries function, and specifies the total number of pixels in the circle-cutout region.

It is to be appreciated that consideration of only some of the four parameters is not sufficient to determine the laparoscopic skills of the trainee. For example, a trainee may perform a cut with a small SD but with a large number of peaks due to inaccurate control of the laparoscopic scissors and grasper and such performance is not considered satisfactory. In a further example, a trainee may perform a cut with a small SD and a small number of peaks but such cut may still be considered unsatisfactory if skewness achieved in such cut is mainly in one direction. The authors believe that a measure of adequate performance may be obtained by a score that takes into consideration performance with respect to a combination of all four parameters as set out hereinafter.

Scoring formula for assessing trainees' performance

The following formula was developed to evaluate the cutting performance of a trainee, i, at every training session, j, in the BTS:

and trainee's circle-cutout, respectively, and S_r and S_t are skewness of reference circle-cutout and trainee's circle-cutout, respectively. NP is the number of peaks and R is the roundness of the circle-cutout (Appendix).

The area represents a measure of the "time" the trainee was outside or inside the perfect circle, and how fast the excursions were corrected. However, the area itself is not a dominant indicator of the trainee's circle-cutout performance. If the area of a trainee's circle-cutout is *larger* than the area of the reference circle ($A_t > A_r$), this could stem from large excursions *outside* the perfect circle shape. However, if the area of a trainee's circle-cutout is *smaller* or *equal* to the reference circle-cutout, this could be due to good performance (perfect roundness of a circle without any excursions) or from zero net of excursions inside and outside the circle shape. Therefore, the relative area was multiplied by the roundness of the circle-cutout.

As indicated from Eq. (1), there are four weights that should be determined. The main question was how to weigh every parameter in Eq. (1), without giving extra points to parameters that are correlated to one another. To meet this goal, we first determined whether the aforementioned variables were correlated. This was done by calculating the Pearson correlation coefficient between every pair of variables [i.e., between number of peaks and standard deviation (NP − SD), relative area and standard deviation (A − SD), number of peaks and relative area (NP − A), skewness and relative area (S − A), skewness and number of peaks (S − NP), and skewness and standard deviation (S − SD)] for all 27 trainees (Appendix). Fig. S1 depicts a heatmap of the Pearson correlation coefficient between six pairs of parameters. As indicated from this heatmap, there is no clear correlation between all 6 pairs because most of the coefficients ranged between − 1 and 1 ($-1 < \rho < 1$).

It is appreciated that although the performance in respect of the combination of the four parameters is ultimately expressed by a score, the data on achievements with respect to each parameter in each training session are also retained to enable the trainee or the instructor (or both) to identify possible deviating patterns and in response to devise ways to improve performance in respect of one or more of the parameters in which the deviating patterns were identified. The retained data

$$Score(i,j) = 100 - \left[w_1 \times \left(1 - \frac{\sigma_r}{\sigma_t} \right) + w_2 \times \frac{|A_r - A_t|}{A_r} \times (1 - R) + w_3 \times \left(1 - \frac{S_r}{|S_t|} \right) \right] - w_4 \times NP \qquad (1)$$

where w_1, w_2, w_3, w_4 are weight factors, σ_r and σ_t are the SD of the reference circle-cutout and the trainee circle-cutout, respectively, A_r and A_t are areas of reference circle-cutout

may be presented as necessary in conventional tables or graphs.

Results

Figure 2 depicts images of the boundaries of a perfect circle-cutout (A–C) and of a random trainee circle-cutout (D–F). The images relate to a single training task. As can be seen in the inset of Fig. 2B, a reference circle was drawn in black on a yellow patch. The black line has 12.6 pixels in width (2 mm). When cutting, excursions that are confined between the inner or outer radii of the reference circle shape are permitted. Moreover, the inner and outer radii remain constant, and the signature of every circle-cutout is measured relative to these radii.

The radius of a perfect circle-cutout, (produced by cutting a circle shape off-BTS with regular scissors), should be constant at every angle [26]. Figure 2A shows a plot of the inner and outer boundaries of the perfect circle-cutout and Fig. 2B shows the 1D signature of the inner and outer boundaries of the perfect circle-cutout. Note that the boundaries of the perfect circle-cutout have small standard deviations (histogram presented in Fig. 2C) with zero excursions outside or inside its reference radii. A trainee who cuts a circle within these boundaries should receive a perfect score of 100.

Figure 2D shows an image of a trainee's actual circle-cutout and Fig. 2E shows the 1D signature function of the inner and outer boundaries of the trainee's actual circle-cutout. Again, the CALS software tracked the boundaries of the circle-cutout (Fig. 2D) and produced its signature (Fig. 2E). Note that there are 3 peaks that extend beyond the outer reference radius, and 8 peaks that are smaller than the inner reference radius.

Furthermore, the standard deviation (histogram presented in Fig. 2F) of the 1D signature (Fig. 2E) is almost 7 times larger than that of the perfect circle-cutout. This suggests that standard deviation is a good indicator for the evaluation of the cutting performance of trainees in FLS.

In order to use the aforementioned parameters for assessing the cutting performance of trainees in BTS properly, we must understand how these parameters relate to each other. As an example of the relation between these parameters, Fig. 3 shows results of five training sessions performed by four trainees. The results in Fig. 3 are presented as graphs of the following: number of peaks; relative area calculated as $A_r - A_t/A_r$, where A_r is the area of the reference circle
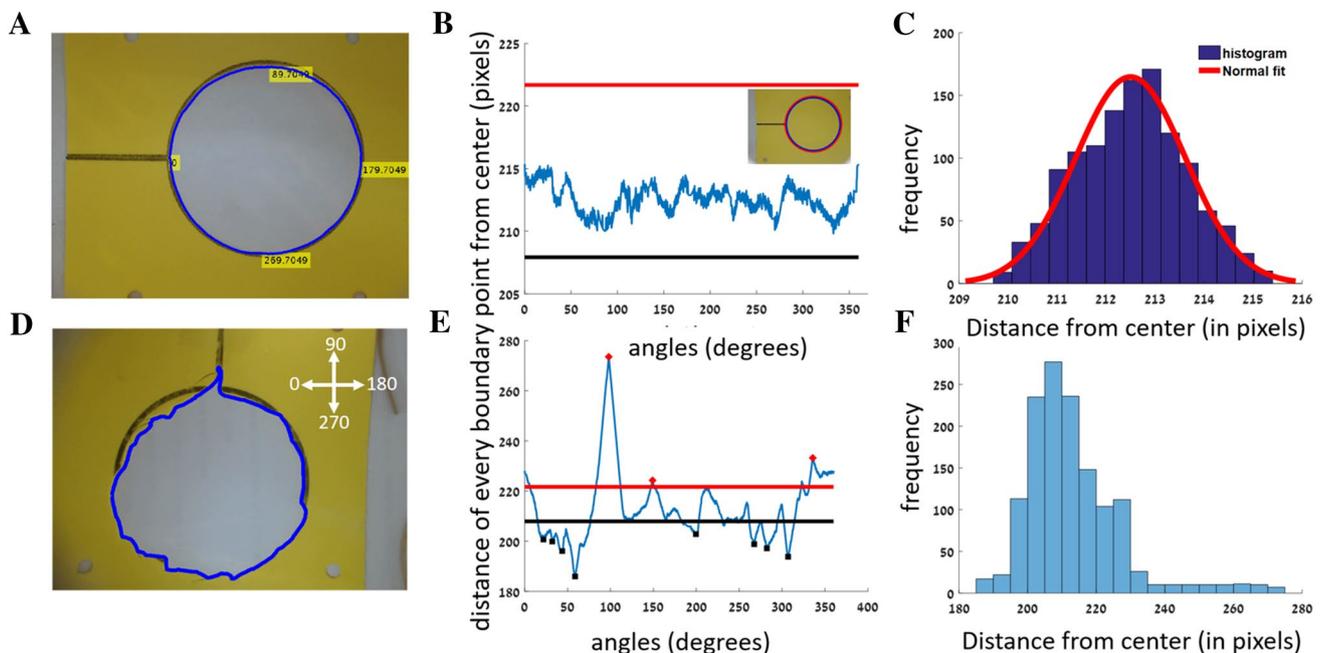

Fig. 2 Boundary extraction from circle-cutout images by CALS. **A** A perfect circle-cutout image, which contains a plot of its inner boundary (blue) circular shape and its outer boundary (red) circular shape. **B** One-dimensional representation (signature) of the perfect circle-cutout boundaries. The red solid line relates to the radius of the outer reference circle (shown in inset), and the black solid line relates to the radius of the inner reference circle. The boundaries of the cutout are confined between the red solid line and the black solid line. **C** Histogram of the perfect circle-cutout radii. Note that the radii have a normal distribution, with standard deviation of 1.03, skewness of 0.036, and kurtosis of 2.42. **D** An example to realistic circle-cutout during laparoscopic training (cut by trainee #11 in training 1), and a plot of the boundaries of the circle-cutout. **E** One-dimensional representation (signature) of the circle-cutout boundaries of **D**. The marked peaks and valleys indicate excursions from the inner and outer reference circles. **F** Histogram of the circle-cutout radii, with standard deviation of 14.07, skewness of 1.632, and kurtosis of 6.88. (Color figure online)
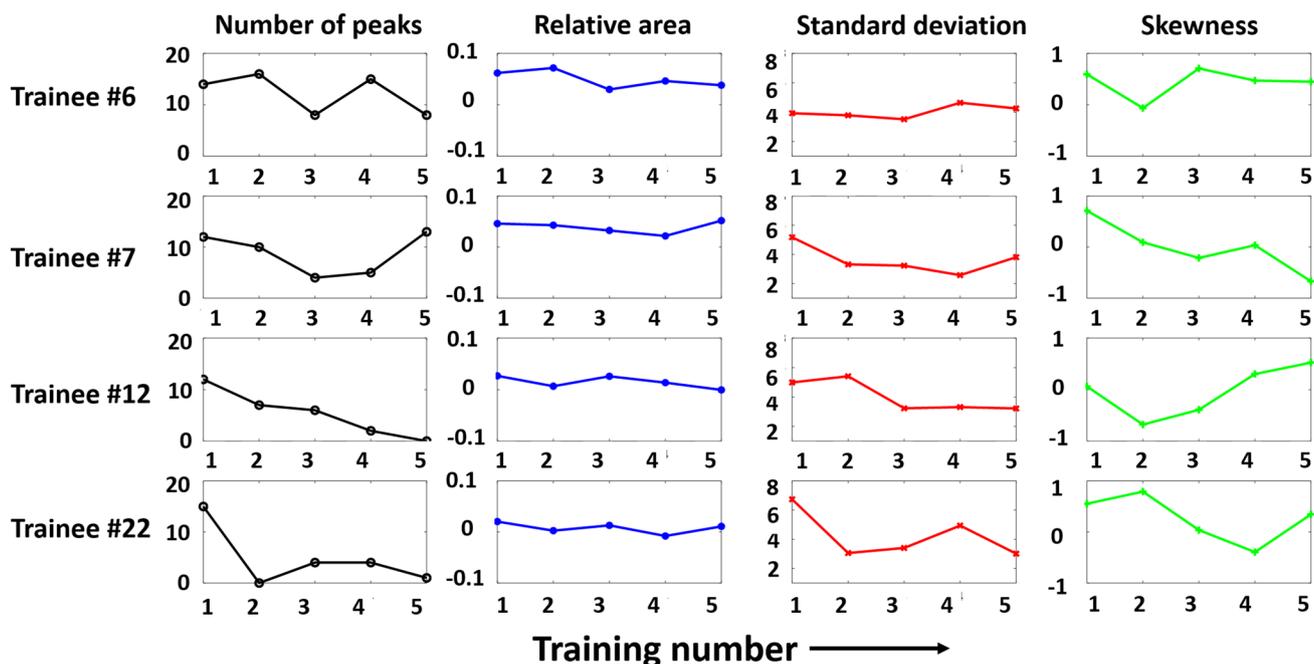
Fig. 3 Graphs of number of peaks, relative area, standard deviation (SD), and skewness achieved by 4 different trainees during 5 training sessions. Each trainee completed the five circle-cutout training sessions at different times. In every training session the total number of peaks, the relative circle area ($A_r - A_t/A_r$), the standard deviation of the circle signature, and the skewness were calculated. Note that although trainees gained experience in cutting the circle shape, their outcome did not necessarily improve, as indicated by the increase in number of peaks (trainee 7) or in standard deviation (trainee 6)

and A_t is the area of the circle made in the training session; SD; and skewness.

In our scoring method, we determined that a large SD of circle-cutout signature will decrease the final score by maximum 8 points ($w_1 = 8$), while large skewness will decrease the final score by maximum 2 points ($w_3 = 2$). This is because the SD of the circle-cutout is a meaningful parameter that better reflects the cutting accuracy than the skewness of the circle-cutout signature. The weight w_4 was determined according to the average *NP* of all trainees (which is ~ 7), where on average it reduces ~ 3 points from the total score ($w_4 = 0.4$). Following determination of these weights, CALS generated scores for all 27 trainees at each training session as seen in Table 1. Any increase in the weights would lower the score, so one could determine in advance the desired score of an average trainee.

Figure 4A depicts a histogram of the average final scores of all 27 trainees, and Fig. 4B demonstrates the distribution of scores as a function of experience and the corresponding error ellipse. The error ellipse was calculated from the principal component analysis (PCA) of the score's covariance matrix [28]. Indeed, the average final score of experienced trainees that had performed at least 70 laparoscopic procedures (85.3) was higher than that of inexperienced trainees that had performed less than 70 laparoscopic procedures (80.4). To summarize our results, we showed that our

software CALS generated a score that expressed the quality of circle-cutout. The score of each trainee in every session resulted from data on deviations from optimal performance with respect to a combination of four parameters (number of peaks, relative area, standard deviation, and skewness) that were extracted from the boundaries of every trainee's circle-cutout image.

Discussion

Usually, it is expected that as the cutting experience increases, a trainee's performance would improve [17, 19]. While it appears that experienced interns gained higher scores than inexperienced interns, we should still be cautious in concluding that there is a clear positive linear relation between score and experience in our study. For example, one could expect that trainee 13, that had already performed 190 laparoscopic surgeries, will receive the highest score among all the trainees, but in practice this did not occur. By looking at Fig. 3, one could observe that there is no clear correlation between the amount of training and performance. For example, trainee 7 reduced the number of excursions from the circle (indicating better control of the laparoscopic scissors and grasper) during the first three sessions, but his/her performance in the

Table 1 Scores of trainees in 4–6 circle-cutout sessions

Trainee number, i	Reported number of laparoscopic surgeries performed	Scores at each session, j						Total average score (i, j)
		1	2	3	4	5	6	
1	35	81	72	76	78	–	–	77
2	40	81	82	82	81	–	–	82
3	12	78	72	78	87	–	–	79
4	10	69	69	81	86	–	–	76
5	15	73	77	84	85	–	–	80
6	40	71	74	82	75	82	–	77
7	10	75	80	86	88	75	–	81
8	10	84	84	77	78	–	–	81
9	10	79	79	89	88	–	–	84
10	20	69	76	72	79	79	–	75
11	30	73	70	65	82	75	–	73
12	40	73	86	82	87	93	–	84
13	190	91	83	86	89	83	–	86
14	150	85	91	86	87	91	–	88
15	–	73	87	88	86	84	–	84
16	100	80	82	78	82	90	–	82
17	20	82	88	87	79	–	–	84
18	–	88	91	89	80	–	–	87
19	100	69	78	86	89	85	–	81
20	100	69	88	87	83	86	89	84
21	100	88	85	88	90	79	–	86
22	20	76	92	90	88	90	–	87
23	0	77	91	88	87	87	90	87
24	–	91	87	86	82	90	–	87
25	–	88	87	86	86	86	–	87
26	70	90	91	92	87	–	–	90
27	–	90	85	85	82	84	–	85

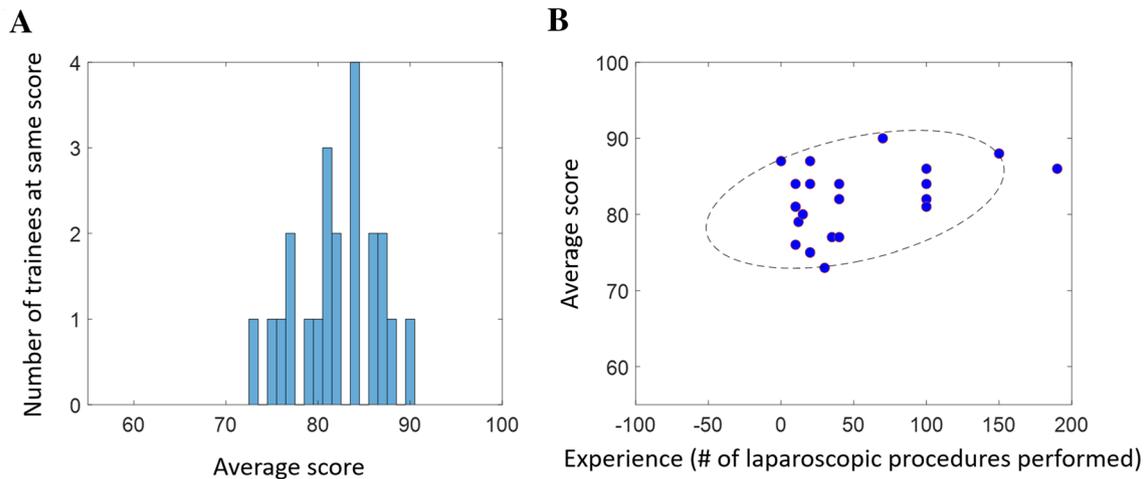Scores are determined by Eq. (1) for $w_1 = 8$, $w_2 = 3000$, $w_3 = 2$, and $w_4 = 0.4$

Fig. 4 **A** Histogram of trainees' scores and relation to their experience (average score 82). **B** Distribution of scores as a function of experience and its correspondence error ellipse

next two sessions, which is expressed as an increase in number of excursions, is poorer although he/she already had three previous training sessions. Furthermore, at first glance it appears that the 4 parameters discussed above are uncorrelated, i.e., there is no clear correlation between relative area, SD, number of excursions and skewness. For example, the SD of trainee 7 has decreased from training session 3–4, but the number of excursions has increased from training session 3–4.

The lack of clear positive correlation between score and accuracy is due to several reasons. The main reason is that no feedback was given to the trainees in response to their scores at every session, and thus a study-curve was absent and trainees did not apply suggestions to improve their scores. In this work, we focused on the diagnostic ability of cutting performance, and less on interpreting the errors, and hence there was no specific training program that referred to the observed deviations and/or mistakes. The instructors that accompanied the trainees did not guide them according to their scores, and all the analysis was done post-training after the last session. Therefore, the trainees that had already completed their training series were not able to post factum correct their mistakes.

Another reason for the lack of clear positive correlation between score and performance is that the time to complete the cutting task was not a factor in our scoring method, in contrast to other studies [19, 23]. Probably, if time was also to be considered a factor (Table S1), different relations between experience and score would have been obtained. For example, experienced trainee 13 completed the cutting task faster than all other trainees. We should however note that time to completion is not a consensus parameter that composes the performance scores of trainees [15]. Following these reasons, further inquiries are needed to conclude whether indeed there is a positive correlation between scores and experience according to our scoring method.

With all this in mind, it is important to note that some improvements in scores from session-to-session can be detected for less experienced trainees (< 15 laparoscopic procedures), similarly to another study [19], even without the instructors' guidance. The fact that there is an improvement in the scores of novice trainees in the BTS as they gain experience in every session makes the trainings and our scoring method favorable especially for inexperienced trainees, and strengthens the need for laparoscopy simulators in hospitals.

In the future, we plan to develop a designated training program that can draw attention to specific cutting inaccuracies (for example, many excursions outside the outer radius) in real time. Then, at a following session, a trainee could be made aware of his past performance to enable him/her to avoid such errors.

Conclusions

In this work we presented an objective scoring method (based on various image-processing algorithms) for assessment of cutting performance in laparoscopic surgical techniques and similar surgical procedures. Our method does not require an instructor for assessment or a self-evaluation questionnaire. For a large set of trainees, every trainee received a score from 0 to 100 that described his/her performance relative to other trainees compatible with the type of internship and professional experience he had accumulated. On average, the total score of experienced interns was higher than the total score of inexperienced interns. However, since the scoring for each training session was given after the trainees had already finished all their cutting tasks and thus were not able to correct their mistakes post factum, and also since no feedback or guidance were given in response to the findings obtained from processing the information, we cannot conclude that there is a definite positive correlation between scores and experience. We also discovered that there was some improvement from session-to-session in the scores of novice trainees, even with lack of feedback from instructors. These findings suggest that our combined BTS-software system could increase the clinical implication of laparoscopic simulators, and eventually better improve surgeons' performances.

Compliance with ethical standards

Appendix

1. The signature of the circle-cutout image is calculated as:

$$r(\theta_i) = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$$

where the centroid points of the circle are calculated as $x_0 = \sum_{i=1}^{m} \sum_{j=1}^{n} iB(i,j) \Big/ \sum_{i=1}^{m} \sum_{j=1}^{n} B(i,j)$, and $y_0 = \sum_{i=1}^{m} \sum_{j=1}^{n} jB(i,j) \Big/ \sum_{i=1}^{m} \sum_{j=1}^{n} B(i,j)$, with $B(i,j)$ equals to 1 inside the circle and 0 outside the circle [27] for circle of size $m \times n$. The centroid of the circle [27] is marked as (x_0, y_0), and every point on its boundary is marked as (x_i, y_i).

2. The Standard Deviation (SD or σ), is defined as:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} \left(r(\theta_i) - \langle r \rangle \right)^2},$$

where N is the number of boundary pixels for the circle-cutout region sampled by the CALS software, $r(\theta_i)$ is defined from Eq. (1), and $\langle r \rangle$ is the mean of the distances $r(\theta_i)$, defined as $\langle r \rangle = 1/N \sum_{i=1}^{N} r(\theta_i)$.

3. The skewness is defined as $S = 1/N \sum_{i=1}^{N} (r(\theta_i) - \langle r \rangle)^3 / \sigma^3$.
4. The area (in pixels) of the circle-cutout specifies the number of pixels in the circle-cutout region, and is defined as $A = \sum_{i=1}^{m} \sum_{j=1}^{n} B(i,j)$, where $B(i,j)$ equals to 1 inside the circle and 0 outside the circle [27].
5. The roundness of the circle-cutout is $R = 4\pi A_t / p^2$, where R is a number from zero to one ($0 \leq R \leq 1$), and p is the circumference of the trainee's circle-cutout, which for the perfect circle equals $2\pi r$, where $r(\theta_i)$ is the radius of the circle. Also for a perfect circle, $A = \pi r^2$ and thus $R = 1$.
6. The rationale behind the formula in Eq. (1) is that a trainee circle-cutout should have SD and skewness close as possible to SD and skewness of a perfect circle $(\sigma_r/\sigma_t \rightarrow 1$ and $S_r/|S_t| \rightarrow 1)$, zero excursions outside and inside a perfect circle shape ($NP = 0$), circle area which is not greater than or lower than the reference circle $(|A_r - A_t|/A_r \rightarrow 0)$, and a roundness close to one, indicating perfectly rounded shape of the circle-cutout ($R \rightarrow 1$). Given these values, the maximum score is 100, since $Score(i,j) = 100 - [w_1 \times (1-1) + w_2 \times 0/A_r \times (1-1) + w_3 \times (1-1)] - w_4 \times 0 = 100$. The minimum score was set to 0, and greatly depends on the weights w_1, w_2, w_3, w_4, since in the "worst case", (i.e., a trainee has huge standard deviation and/or area and/or skewness and/or number of excursion, where $\sigma_r \ll \sigma_t$ and then $\sigma_r/\sigma_t \rightarrow 0$, and/or $A_r \ll A_t$, and/or $S_r \ll S_t$ and/or $NP \gg 1$ and/or $R = 0$) the minimum score will be equal to $Score(i,j) = 100 - w_1 - w_2 \times |A_r - A_t|/(A_r) - w_3 - w_4 \times NP$. Remember that the number of peaks (NP) is the sum of excursions inside and outside the circle-cutout. The average NP of all trainees was ~ 7 excursions. As the NP increases, the score decreases.
7. Pearson correlation coefficient, ρ, is defined as:

$$\rho = \frac{1}{L-1} \sum_{i=1}^{L} \left(\frac{x_i - \mu_x}{\sigma_x} \right) \left(\frac{y_i - \mu_y}{\sigma_y} \right),$$

where L is the sample size (4–6 trainings), x, y are two parameter pairs, μ_x, μ_y and σ_x, σ_y are the mean and SD of the two parameters.

References

1. Preshaw J, Siassakos D, James M, Draycott T, Vyas S, Burden C (2018) Patients and hospital managers want laparoscopic simulation training to become mandatory before live operating: a multicentre qualitative study of stakeholder perceptions. BMJ Simul. https://doi.org/10.1136/bmjstel-2017-000270
2. Csikesz NG, Singla A, Murphy MM, Tseng JF, Shah SA (2010) Surgeon volume metrics in laparoscopic cholecystectomy. Dig Dis Sci 55(8):2398–2405
3. Zitsman JL (2006) Pediatric minimal-access surgery: update. Pediatrics 118(1):304–308
4. Sooriakumaran P, Srivastava A, Shariat SF, Stricker PD, Ahlering T, Eden CG, Wiklund PN, Sanchez-Salas R, Mottrie A, Lee D, Neal DE, Ghavamian R, Nyirady P, Nilsson A, Carlsson S, Xylinas E, Loidl W, Seitz C, Schramek P, Roehrborn C, Cathelineau X, Skarecky D, Shaw G, Warren A, Delprado WJ, Haynes AM, Steyerberg E, Roobol MJ, Tewari AK (2014) A multinational, multi-institutional study comparing positive surgical margin rates among 22393 open, laparoscopic, and robot-assisted radical prostatectomy patients. Eur Urol 66(3):450–456
5. Hidlebaugh DA, Vulgaropulos S, Orr RK (1997) Treating adnexal masses. Operative laparoscopy vs. laparotomy. J Reprod Med 42(9):551–558
6. Magrina JF (2002) Complications of laparoscopic surgery. Clin Obstet Gynecol 45(2):469–480
7. Trottier DC, Martel G, Boushey RP (2009) Complications in laparoscopic intestinal surgery: prevention and management. Minerva Chir 64:339–354
8. Munz Y, Kumar BD, Moorthy K, Bann S, Darzi A (2004) Laparoscopic virtual reality and box trainers: is one superior to the other? Surg Endosc 18:485–494
9. Paquette J, Lemyre M, Vachon-Marceau C, Bujold E, Maheux-Lacroix S (2017) Virtual laparoscopy simulation: a promising pedagogic tool in gynecology. JSLS. https://doi.org/10.4293/JSLS.2017.00048
10. van Empel PJ, van der Veer WM, van Rijssen LB, Cuesta MA, Scheele F, Bonjer HJ, Meijerink WJ (2010) Mapping the maze of minimally invasive surgery simulators. J Lap Adv Surg Tech A 22:51–60
11. Peters JH, Fried GM, Swanstrom LL, Soper NJ, Sillin LF, Schirmer B, Hoffman K, SAGES FLS Committee (2004) Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. Surgery 135:21–27
12. Shiffer C (2008) ABS to require ACLS, ATLS and FLS for general surgery certification. https://www.absurgery.org/default.jsp?news_newreqs
13. Scott DJ, Ritter EM, Tesfay ST, Pimentel EA, Nagji A, Fried GM (2008) Certification pass rate of 100% for fundamentals of laparoscopic surgery skills after proficiency-based training. Surg Endosc 22:1887
14. Sroka G, Feldman LS, Vassiliou MC, Kaneva PA, Fayez R, Fried GM (2010) Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room—a randomized controlled trial. AM J Surg 199(1):115–120
15. Beyer L, Troyer JD, Mancini J, Bladou F, Berdah SV, Karsenty G (2011) Impact of laparoscopy simulator training on the technical skills of future surgeons in the operating room: a prospective study. AM J Surg 202(3):265–272
16. De Win G, Van Bruwaene S, Kulkarni J, Van Calster B, Aggarwal R, Allen C, Lissens A, De Ridder D, Miserez M (2016) An evidence-based laparoscopic simulation curriculum shortens the

clinical learning curve and reduces surgical adverse events. Adv Med Edu Prac 7:357–370

17. McCluney AL, Vassiliou MC, Kaneva PA, Cao J, Stanbridge DD, Feldman LS, Fried GM (2007) FLS simulator performance predicts intraoperative laparoscopic skill. Surg Endosc 21(11):1991–1995

18. Wong J, Bhattacharya G, Vance SJ, Bistolarides P, Merchant AM (2013) Construction and validation of a low-cost laparoscopic simulator for surgical education. J Surg Educ 70(4):443–450

19. Shepherd G, von Delft D, Truck J, Kubiak R, Ashour K, Grant H (2016) A simple scoring system to train surgeons in basic laparoscopic skills. Pediatr Surg Int 32(3):245–252

20. Nataraja RM, Ade-Ajayi N, Curry JI (2006) Surgical skills training in the laparoscopic era: the use of a helping hand. Pediatr Surg Int 22:1015–1020

21. Van Hove PD, Tuijthof GJ, Verdaasdonk EG, Stassen LP, Dankelman J (2010) Objective assessment of technical surgical skills. Br J Surg 97:972–987

22. Brackmann MW, Andreatta P, McLean K, Reynolds RK (2016) Development of a novel simulation model for assessment of laparoscopic camera navigation. Surg Endosc. 31:3033–3039

23. Chmarra MK, Klein S, de Winter JCF, Jansen FW, Dankelman J (2010) Objective classification of residents based on their psychomotor laparoscopic skills. Surg Endosc 24(5):1031–1039

24. Ghasemloonia A, Maddahi Y, Zareinia K, Lama S, Dort JC, Sutherland GR (2017) Surgical skill assessment using motion quality and smoothness. J Surg Educ 74(2):295–305

25. Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondré K, Stanbridge D, Fried GM (2005) A global assessment tool for evaluation of intraoperative laparoscopic skills. Am J Surg 190:107–113

26. Gonzalez RC (2007) Digital image processing, 3rd edn. Pearson International Edition, New Delhi

27. Jain R, Kasturi R, Schunck BG (1995) Machine Vision. McGraw-Hill, New York

28. Jolliffe IT (2002) Principal component analysis, 2nd edn. Springer, New York