**2022 SAGES ORAL**

# Artificial intelligence in laparoscopic simulation: a promising future for large-scale automated evaluations

Francisca Belmar[1] · María Inés Gaete[1] · Gabriel Escalona[1] · Martín Carnier[1] · Valentina Durán[1] · Ignacio Villagrán[1,2] · Domenech Asbun[3] · Matías Cortés[1] · Andrés Neyem[4] · Fernando Crovari[1,2] · Adnan Alseidi[5] · Julián Varas[1,2]

## Abstract

**Introduction** A limitation to expanding laparoscopic simulation training programs is the scarcity of expert evaluators. In 2019, a new digital platform for remote and asynchronous laparoscopic simulation training was validated. Through this platform, 369 trainees have been trained in 14 institutions across Latin America, collecting 6729 videos of laparoscopic training exercises. The use of artificial intelligence (AI) has recently emerged in surgical simulation, showing usefulness in training assessment, virtual reality scenarios, and laparoscopic virtual reality simulation. An AI algorithm to assess basic laparoscopic simulation training exercises was developed. This study aimed to analyze the agreement between this AI algorithm and expert evaluators in assessing basic laparoscopic-simulated training exercises.

**Methods** The AI algorithm was trained using 400-bean drop (BD) and 480-peg transfer (PT) videos and tested using 64-BD and 43-PT randomly selected videos, not previously used to train the algorithm. The agreement between AI and expert evaluators from the digital platform (EE) was then analyzed. The exercises being assessed involve using laparoscopic graspers to move objects across an acrylic board without dropping any objects in a determined time (BD < 24 s, PT < 55 s). The AI algorithm can detect object movement, identify if objects have fallen, track grasper clamps location, and measure exercise time. Cohen's Kappa test was used to evaluate the agreement between AI assessments and those performed by EE, using a pass/fail nomenclature based on the time to complete the exercise.

**Results** After the algorithm was trained, 79.69% and 93.02% agreement were observed in BD and PT, respectively. The Kappa coefficients test observed for BD and PT were 0.59 (moderate agreement) and 0.86 (almost perfect agreement), respectively.

**Conclusion** This first approach of AI use in basic laparoscopic skills simulated training assessment shows promising results, providing a preliminary framework to expand the use of AI to other basic laparoscopic skills exercises.

**Keywords** Artificial intelligence · Surgical simulation · Laparoscopy

✉ Julián Varas
jevaras@uc.cl

1 Experimental Surgery and Simulation Center, Department of Digestive Surgery, Catholic University of Chile, Santiago, Chile

2 Faculty of Medicine, Catholic University of Chile, Santiago, Chile

3 Hepatobiliary & Pancreatic Surgery, Miami Cancer Institute, Miami, USA

4 Department of Computer Science, Engineering Faculty, Catholic University of Chile, Santiago, Chile

5 Department of Surgery, University of California San Francisco, San Francisco, USA

## Laparoscopic simulation training

Laparoscopic surgery has become the gold standard approach for several procedures due to its lower postoperative pain, infection rates, and hospital stay [1–4]. Therefore, it is mandatory for surgeons to achieve good laparoscopic skills. This can be achieved with safe practice through simulation training, which has proven to be effective in the acquisition and transfer of technical skills to the operating room [5, 6]. Basic laparoscopic simulation training programs are widely used in surgical education with multiple programs available [5–11]. The most widely used laparoscopic skills training program is the Fundamentals of Laparoscopic Surgery (FLS), which was developed by

the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) and released to the public in 2004 [7, 12]. In 2008, successful completion of the FLS program became a requirement for board certification by the American Board of Surgery, reflecting its importance in surgical education [13].

Multiple advanced laparoscopic training programs have been created. In 2012, an advanced laparoscopic skills curriculum was developed at our institution, based on the creation of an enteric anastomosis using ex vivo bovine intestine [5]. This program has shown marked improvement of technical skills in a simulated scenario and also demonstrated successful transfer of these skills to the operating room. Trainees who underwent the program attained a proficiency level similar to practicing laparoscopic surgical experts [5, 6].

In 2019, the minimally invasive telementoring opportunity (MITO) project was developed to provide advanced laparoscopic skills training in remote places without expert evaluators immediately available [14]. In this study, a digital platform called "LAPP" was created to allow for remote administration of the previously mentioned validated advanced laparoscopic skills training program [5], which was possible through continuous feedback given in a remote and asynchronous manner. It was found that the group trained remotely through the platform acquired laparoscopic skills comparable to those trained through the same program with in-person direction and feedback. The good results achieved with this platform allowed the expansion of this training program to multiple sites and allowed continuous practice even through the COVID-19 pandemic [15]. Currently, this training program is available in fourteen cities across eight countries and has reached over 350 trainees in less than 5 years. The development of this new training system has allowed the storage of over 6500 videos of laparoscopic skills training.

## Artificial intelligence uses in surgery

The incorporation of AI in surgery includes the creation of algorithms capable of pattern recognition. The use of AI has been observed in surgical simulation for the assessment of training, virtual reality scenarios, and laparoscopy virtual reality simulation [16–18]. For example, AI algorithms are currently under development for the assessment of suturing skills in virtual reality scenarios for robotic surgery [19]. In addition, it has been applied to clinical scenarios, specifically in robotic and laparoscopic surgery through the detection of anatomy in the surgical field, aiming to be applied as a guideline for these procedures [18–22].

Despite the success of the MITO project and other validated and scalable training programs [5–7, 14], the finite number of evaluators available to assess trainees remains a limiting factor. It is in this context that the incorporation of AI can be used in a favorable way to create a solution

that allows for the mass evaluation of trainees using digital platforms. Therefore, through this study, we present innovation in mass trainee assessment that incorporates machine learning and other methodologies (such as deep learning) to create automatic assessment algorithms. Our aim was to develop an AI algorithm capable of evaluating basic laparoscopic skills training exercises with results similar to those of expert evaluators.

## Methods

### Laparoscopic skills exercises

A previously developed and validated basic laparoscopic skills training program is available through a digital platform and has been taught to 369 trainees. Currently, the platform database gathers 6729 videos of trainees performing basic and advanced laparoscopic skills training exercises. Data related to each exercise was extracted and analyzed by expert data science engineers. This included all videos of basic laparoscopic exercises performed by trainees and uploaded to the digital platform. This consisted of a total of 6496 video-recorded training of 11 basic laparoscopic exercises.

After data organization, two exercises were selected for the training of the machine learning algorithm. These were the bean drop (BD) and peg transfer (PT) exercises. The BD exercise consists of moving five beans using laparoscopic graspers from one box to another. The goal is to avoid dropping any beans and complete the exercise in less than 24 s (Fig. 1a). The PT exercise involves transferring six rubber objects from one side of a pegboard to the contralateral side, and back, using laparoscopic graspers. This exercise must be performed in less than 55 s while handing the objects from one grasper to the other in mid-air without dropping any of them (Fig. 1b).

### Algorithm development

An AI algorithm was developed using a Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology as a scientific computing framework using Python, and Pytorch [23, 24] as a framework for convolutional neural networks. U-net is used for segmentation, allowing clamp tracking, and YOLO v4 for element, receptacle, and pin detection. For the labeling process, an expert group of AI labelers was required to label random video frames of the exercises and provided Pascal VOC files.

This algorithm uses fragmented video photograms and Pascal VOC files to detect the position of the grasper clamps within the working environment and object movement while identifying if objects have fallen (Fig. 2). To develop the model, all videos available on the platform from both exercises until March 2021 were included for training, this consisted of 400 BD and 480

**Fig. 1** Exercise examples extracted from the platform database. On the left: **a** Bean drop and on the right side: **b** Peg transfer
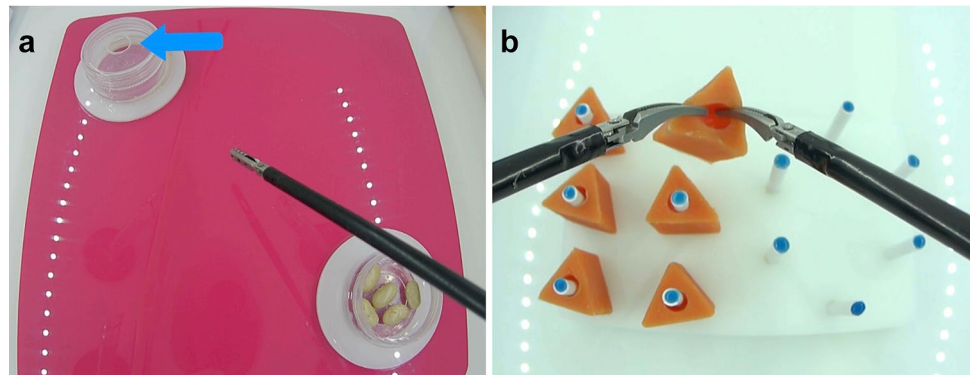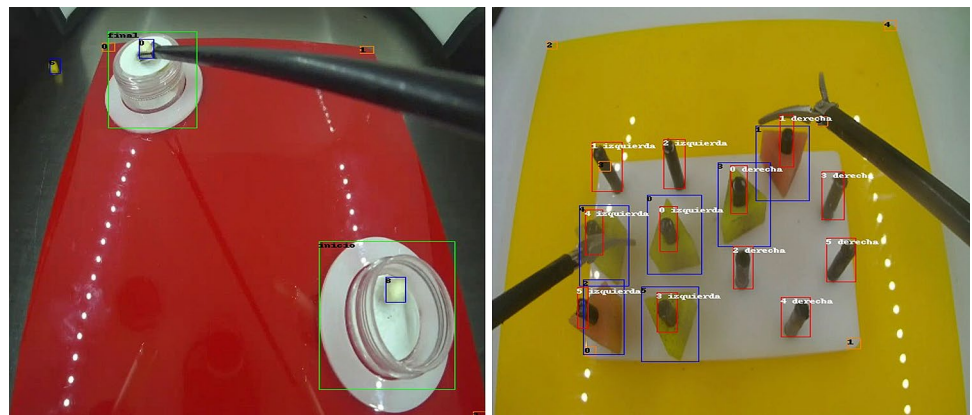


**Fig. 2** Object detection observed with the algorithm. Bean drop (left) and peg transfer (right)

PT videos; then 64 videos of BD and 43-PT videos were used for testing. The algorithm provides two main outputs: (1) falling objects and (2) time to complete the exercise. The falling of objects was defined as the objects not being in contact with the graspers, the peg board, or the receptacles. Time measurements include time from first to last contact between the graspers and the objects, and these measurements were manually categorized as pass or fail according to the previously established passing times and then compared to the actual gold standard, expert evaluators.

## Statistical analysis

Data were analyzed with RStudio [25] using Cohen's Kappa test to assess agreement between AI and expert evaluators.

## Results

The developed algorithm U-net segmentation achieved a 98% precision for the location of the grasper clamps in the video frame. Afterward, the platform was tested for both of the previously mentioned exercises.

A high level of agreement was observed between the AI algorithm and expert evaluators for the peg transfer exercise, with a 93.02% of agreement. The observed Cohen's Kappa coefficient was 0.86 showing an almost perfect agreement (Table 1).

Meanwhile, for the bean drop, a 79.69% agreement with expert evaluators was observed, with a Cohen's Kappa coefficient of 0.59, which means the algorithm has a moderate agreement with the current gold standard (Table 2).

## Discussion

The results from this study show that it is not only feasible to develop AI algorithms to assess basic laparoscopic simulation training exercises but also the application of AI can

**Table 1** Summary of the assessments of peg transfer exercises performed by artificial intelligence (AI) and expert evaluators (EE)

|  | EE | |
| --- | --- | --- |
|  | Pass | Fail |
| AI |  |  |
| Pass | 19 | 2 |
| Fail | 1 | 21 |

**Table 2** Summary of the assessments of bean drop exercises performed by artificial intelligence (AI) and expert evaluators (EE)

|  | EE | |
|---|---|---|
|  | Pass | Fail |
| AI |  |  |
| Pass | 29 | 4 |
| Fail | 9 | 22 |

have high levels of agreement with the current gold standard (expert evaluators).

Currently, our engineering and programming teams are working on developing AI algorithms for all eleven basic laparoscopic-simulated training exercises available through the digital platform.

If AI development, and further application, is successful for assessment through this digital platform, it could also be applied to other similar basic laparoscopic skills training programs. Therefore, the certification of basic laparoscopic skills could be provided from anywhere in the world, without the need for expert evaluators available to perform the assessment. This would make it easier for trainees to achieve their proficiency certification, without needing to travel long distances to simulation centers with expert evaluators, or the need for evaluators to be synchronously assessing through video conferencing platforms.

Although the findings presented in this study are promising, it is important to mention that they are not completely free of limitations. First, the videos used for calibration of the algorithm and testing were all part of a standardized training program; therefore, it is unknown if these algorithms can be applied to other training programs with similar results.

Secondly, the algorithm was developed using less than 500 sample videos for each exercise, achieving good agreement with the gold standard. If more videos are collected, then the algorithm could be modified to improve the accuracy even further.

Third, the current AI algorithm has technical limitations in labeling fallen objects, especially in the bean drop exercise, where it is more frequent. For this reason, although it can be obtained as an output of the algorithm it was not considered among the pass or fail criteria; however, it is currently being improved to increase the accuracy of the algorithm.

Additionally, to avoid manually categorizing the exercises as pass or fail based on the algorithm outputs, modifications to the digital platform are being developed to automatically incorporate the algorithm outputs.

Finally, it is equally important to emphasize that, for now, this algorithm can only measure the time it takes to complete the exercise, and it does not replace experts' feedback.

To summarize, even though the AI algorithm developed provides simple outputs, we believe the results observed through this study are promising for the automated assessment of basic simulated laparoscopic skills and can be expanded to more exercises.

## Declarations

## References

1. Li Y, Xiang Y, Wu N, Wu L, Yu Z, Zhang M, Wang M, Jiang J, Li Y (2015) A comparison of laparoscopy and laparotomy for the management of abdominal trauma: a systematic review and meta-analysis. World J Surg 39:2862–2871. https://doi.org/10.1007/s00268-015-3212-4

2. Hutter MM, Randall S, Khuri SF, Henderson WG, Abbott WM, Warshaw AL (2006) Laparoscopic versus open gastric bypass for morbid obesity: a multicenter, prospective, risk-adjusted analysis from the National Surgical Quality Improvement Program. Ann Surg 243:657–662. https://doi.org/10.1097/01.sla.0000216784.05951.0b (**discussion 662–666**)

3. Sauerland S, Walgenbach M, Habermalz B, Seiler CM, Miserez M (2011) Laparoscopic versus open surgical techniques for ventral or incisional hernia repair. Cochrane Database Syst Rev. https://doi.org/10.1002/14651858.CD007781.pub2

4. Velanovich V (2000) Laparoscopic vs open surgery: a preliminary comparison of quality-of-life outcomes. Surg Endosc 14:16–21. https://doi.org/10.1007/s004649900003

5. Varas J, Mejía R, Riquelme A, Maluenda F, Buckel E, Salinas J, Martínez J, Aggarwal R, Jarufe N, Boza C (2012) Significant transfer of surgical skills obtained with an advanced laparoscopic training program to a laparoscopic jejunojejunostomy in a live porcine model: feasibility of learning advanced laparoscopy in a general surgery residency. Surg Endosc 26:3486–3494. https://doi.org/10.1007/s00464-012-2391-4

6. Boza C, León F, Buckel E, Riquelme A, Crovari F, Martínez J, Aggarwal R, Grantcharov T, Jarufe N, Varas J (2017) Simulation-trained junior residents perform better than general surgeons on

advanced laparoscopic cases. Surg Endosc 31:135–141. https://doi.org/10.1007/s00464-016-4942-6

7. Peters JH, Fried GM, Swanstrom LL, Soper NJ, Sillin LF, Schirmer B, Hoffman K, the 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. https://doi.org/10.1016/S0039-6060(03)00156-9

8. León Ferrufino F, Varas Cohen J, Buckel Schaffner E, Crovari Eulufi F, Pimentel Müller F, Martínez Castillo J, Jarufe Cassis N, Boza Wilson C (2015) Simulation in laparoscopic surgery. Cirugía Española (English Edition) 93:4–11. https://doi.org/10.1016/j.cireng.2014.02.022

9. Supe A, Prabhu R, Harris I, Downing S, Tekian A (2012) Structured training on box trainers for first year surgical residents: does it improve retention of laparoscopic skills? A randomized controlled study. J Surg Educ 69:624–632. https://doi.org/10.1016/j.jsurg.2012.05.002

10. Boza C, Varas J, Buckel E, Achurra P, Devaud N, Lewis T, Aggarwal R (2013) A cadaveric porcine model for assessment in laparoscopic bariatric surgery—a validation study. Obes Surg 23:589–593. https://doi.org/10.1007/s11695-012-0807-9

11. Charokar K, Modi JN (2021) Simulation-based structured training for developing laparoscopy skills in general surgery and obstetrics & gynecology postgraduates. J Educ Health Promot 10:387. https://doi.org/10.4103/jehp.jehp_48_21

12. Fundamentals of Laparoscopic Surgery. https://www.flsprogram.org/. Accessed 2 Mar 2022

13. Cullinan DR, Schill MR, DeClue A, Salles A, Wise PE, Awad MM (2017) Fundamentals of laparoscopic surgery: not only for senior residents. J Surg Educ 74:e51–e54. https://doi.org/10.1016/j.jsurg.2017.07.017

14. Quezada J, Achurra P, Jarry C, Asbun D, Tejos R, Inzunza M, Ulloa G, Neyem A, Martínez C, Marino C, Escalona G, Varas J (2020) Minimally invasive tele-mentoring opportunity—the mito project. Surg Endosc 34:2585–2592. https://doi.org/10.1007/s00464-019-07024-1

15. Jarry Trujillo C, Achurra Tirado P, Escalona Vivas G, Crovari Eulufi F, Varas Cohen J (2020) Surgical training during COVID-19: a validated solution to keep on practicing. Br J Surg 107:e468–e469. https://doi.org/10.1002/bjs.11923

16. Barua I, Vinsard DG, Jodal HC, Løberg M, Kalager M, Holme Ø, Misawa M, Bretthauer M, Mori Y (2021) Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. Endoscopy 53:277–284. https://doi.org/10.1055/a-1201-7165

17. Alaker M, Wynn GR, Arulampalam T (2016) Virtual reality training in laparoscopic surgery: a systematic review & meta-analysis. Int J Surg 29:85–94. https://doi.org/10.1016/j.ijsu.2016.03.034

18. Vedula SS, Ghazi A, Collins JW, Pugh C, Stefanidis D, Meireles O, Hung AJ, Schwaitzberg S, Levy JS, Sachdeva AK, the Collaborative for Advanced Assessment of Robotic Surgical Skills (2022) Artificial intelligence methods and artificial intelligence-enabled metrics for surgical education: a multidisciplinary consensus. J Am Coll Surg 234:1181–1192. https://doi.org/10.1097/XCS.0000000000000190

19. Hung AJ, Rambhatla S, Sanford DI, Pachauri N, Vanstrum E, Nguyen JH, Liu Y (2021) Road to automating robotic suturing skills assessment: battling mislabeling of the ground truth. Surgery. https://doi.org/10.1016/j.surg.2021.08.014

20. Tokuyasu T, Iwashita Y, Matsunobu Y, Kamiyama T, Ishikake M, Sakaguchi S, Ebe K, Tada K, Endo Y, Etoh T, Nakashima M, Inomata M (2021) Development of an artificial intelligence system using deep learning to indicate anatomical landmarks during laparoscopic cholecystectomy. Surg Endosc 35:1651–1658. https://doi.org/10.1007/s00464-020-07548-x

21. Madani A, Namazi B, Altieri MS, Hashimoto DA, Rivera AM, Pucher PH, Navarrete-Welton A, Sankaranarayanan G, Brunt LM, Okrainec A, Alseidi A (2020) Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. Ann Surg. https://doi.org/10.1097/SLA.0000000000004594

22. Moglia A, Ferrari V, Morelli L, Ferrari M, Mosca F, Cuschieri A (2016) A systematic review of virtual reality simulators for robot-assisted surgery. Eur Urol 69:1065–1080. https://doi.org/10.1016/j.eururo.2015.09.021

23. Van Rossum G, Drake Jr FL (1995) Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam

24. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A, Kopf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J, Chintala S (2019) PyTorch: an imperative style, high-performance deep learning library. In: Advances in neural information processing systems. Curran Associates Inc, Red Hook

25. RStudio | Open source & professional software for data science teams. https://www.rstudio.com/. Accessed 4 Mar 2022