

## REVIEW ARTICLE OPEN



# Computer vision in surgery: from potential to clinical value

Pietro Mascagni<sup>1,2,3,12</sup>, Deepak Alapatt<sup>4,12</sup>, Luca Sestini<sup>4,5</sup>, Maria S. Altieri<sup>3,6</sup>, Amin Madani<sup>3,7</sup>, Yusuke Watanabe<sup>3,8</sup>, Adnan Alseidi<sup>3,9</sup>, Jay A. Redan<sup>10</sup>, Sergio Alfieri<sup>11</sup>, Guido Costamagna<sup>11</sup>, Ivo Boškoski<sup>11</sup>, Nicolas Padoy<sup>2,4</sup> and Daniel A. Hashimoto<sup>3,6</sup>

Hundreds of millions of operations are performed worldwide each year, and the rising uptake in minimally invasive surgery has enabled fiber optic cameras and robots to become both important tools to conduct surgery and sensors from which to capture information about surgery. Computer vision (CV), the application of algorithms to analyze and interpret visual data, has become a critical technology through which to study the intraoperative phase of care with the goals of augmenting surgeons' decision-making processes, supporting safer surgery, and expanding access to surgical care. While much work has been performed on potential use cases, there are currently no CV tools widely used for diagnostic or therapeutic applications in surgery. Using laparoscopic cholecystectomy as an example, we reviewed current CV techniques that have been applied to minimally invasive surgery and their clinical applications. Finally, we discuss the challenges and obstacles that remain to be overcome for broader implementation and adoption of CV in surgery.

*npj Digital Medicine* (2022)5:163; <https://doi.org/10.1038/s41746-022-00707-5>

With over 330 million procedures performed annually, surgery represents a critical segment of healthcare systems worldwide<sup>1</sup>. Surgery, however, is not readily accessible to all. The Lancet Commission on Global Surgery estimated that 143 million additional surgical procedures are needed each year to “save lives and prevent disability”<sup>2</sup>. Improvements in perioperative care and the introduction of minimally invasive approaches have made the surgery more effective but also more complex and expensive, with surgery accounting for about one-third of U.S. healthcare costs<sup>3</sup>. Furthermore, a large proportion of preventable medical errors happen in operating rooms (OR)<sup>4</sup>. These observations suggest the need for developing solutions to improve surgical safety and efficiency.

The analysis of videos of surgical procedures and OR activities could offer strategies to improve this critical phase of surgical care. This is especially true for procedures performed with a minimally invasive approach, which is being increasingly adopted globally<sup>5–7</sup> and heavily relies on the visualization provided by fiber optic cameras. In fact, in minimally invasive surgery the partial loss of haptic feedback is compensated by magnified, high-definition videos acquired by endoscopic cameras<sup>8</sup>. Endoscopic videos guiding surgical procedures represent a direct and readily available source of digital data on the intraoperative phase of surgical care.

In recent years, the analysis of endoscopic videos of minimally invasive surgical procedures has enabled the study of the impact of OR activities on patient outcomes<sup>9</sup> and the assessment of quality improvement initiatives<sup>10</sup>. In addition, video-based assessment (VBA) is being increasingly investigated for operative performance assessment, formative feedback, and surgical credentialing. However, VBA has mostly remained confined to the research domain given the burden of manually reviewing and consistently assessing surgical videos<sup>11,12</sup>. Expanding on initial

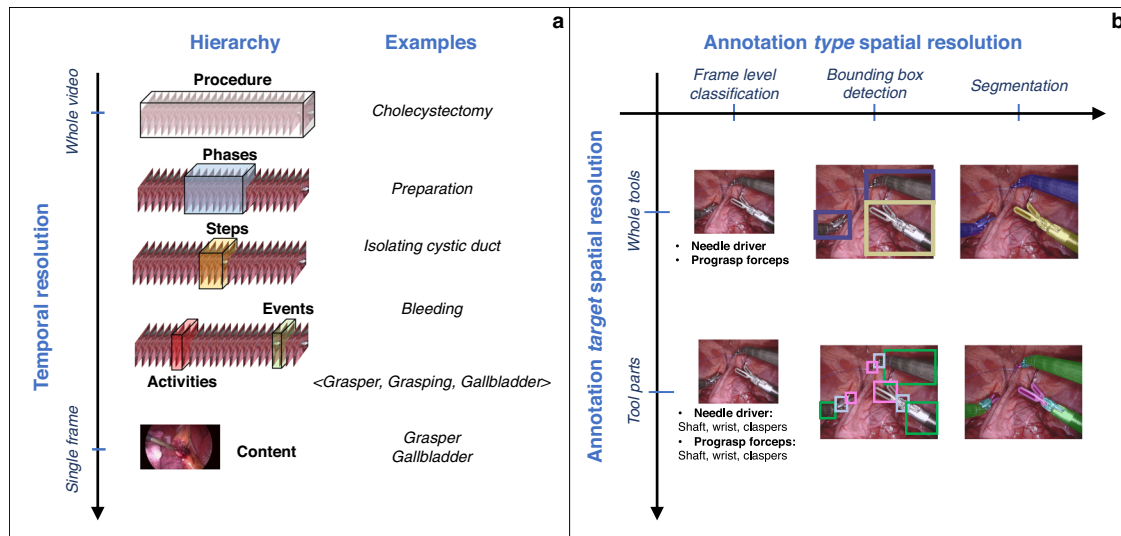
successes in minimally invasive surgery, use of video has been growing in open surgery as well<sup>13</sup>.

Computer vision (CV), a computer science discipline that utilizes artificial intelligence (AI) techniques such as deep learning (DL) to process and analyze visual data, could facilitate endoscopic video analysis and allow scaling of applications for the benefit of a wider group of surgeons and patients<sup>14</sup>. Furthermore, while humans tend to grossly assess images qualitatively, computer algorithms have the potential to extract invisible, *quantitative*, and objective information on intraoperative events. Finally, automated, online, endoscopic video analysis could allow us to monitor cases in real-time, predict complications, and intervene to improve care and prevent adverse events.

Recently, several DL-based CV solutions mostly for minimally invasive surgery have been developed by academics as well as industry groups. CV applications range from workflow analysis to automated performance assessment. While analogous digital solutions are being clinically translated and implemented at scale for diagnostic applications in gastrointestinal endoscopy<sup>15</sup> and radiology<sup>16</sup>, CV in surgery is lagging.

We discuss the current state, potential, and possible paths toward the clinical value of computer vision in surgery. We examined laparoscopic cholecystectomy, currently the most studied surgical procedure for CV methods, to provide a specific example of how CV has been approached in surgery; however, many of these methods have been applied to robotic, endoscopic, and open surgery as well. Finally, we discuss recent efforts to improve access and methods to better model surgical data together with the ethical, legal, and educational considerations fundamental to delivering value to patients, clinicians, and healthcare systems.

<sup>1</sup>Gemelli Hospital, Catholic University of the Sacred Heart, Rome, Italy. <sup>2</sup>IHU-Strasbourg, Institute of Image-Guided Surgery, Strasbourg, France. <sup>3</sup>Global Surgical Artificial Intelligence Collaborative, Toronto, ON, Canada. <sup>4</sup>ICube, University of Strasbourg, CNRS, IHU, Strasbourg, France. <sup>5</sup>Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milano, Italy. <sup>6</sup>Department of Surgery, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. <sup>7</sup>Department of Surgery, University Health Network, Toronto, ON, Canada. <sup>8</sup>Department of Surgery, University of Hokkaido, Hokkaido, Japan. <sup>9</sup>Department of Surgery, University of California San Francisco, San Francisco, CA, USA. <sup>10</sup>Department of Surgery, AdventHealth-Celebration Health, Celebration, FL, USA. <sup>11</sup>Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy. <sup>12</sup>These authors contributed equally: Pietro Mascagni, Deepak Alapatt. ✉email: [pietro.mascagni@ihu-strasbourg.eu](mailto:pietro.mascagni@ihu-strasbourg.eu)



**Fig. 1 Framework for the analysis of endoscopic videos.** Temporal (a) and spatial (b) annotations at different resolutions are used to model tasks at increasingly finer details.

### COMPUTER VISION FOR LAPAROSCOPIC CHOLECYSTECTOMY

Cholecystectomy is the most common abdominal surgical procedure, with almost one million cases performed in the US alone each year<sup>17</sup>. The safety and efficacy of minimally invasive surgery were demonstrated over two decades ago, and laparoscopy has since become the gold standard approach for the removal of the gallbladder. Laparoscopic cholecystectomy (LC) generally follows a standardized operative course, is performed by most general surgeons, and is often one of the first procedures introduced during surgical training. A relatively recent analysis pooling data from more than five thousand patients confirmed the safety of LC, reporting 1.6–5.3% and 0.08–0.14% overall morbidity and mortality rates, respectively<sup>17</sup>. Nonetheless, iatrogenic bile duct injuries (BDIs) still complicate 0.32–1.5% of LCs<sup>17,18</sup>, rates higher than the incidence commonly reported in open surgery<sup>19</sup>. BDIs resulted in a three-fold increase in mortality at one year, a lifelong decrease in quality of life despite expert repair, and were estimated to have an annual cost of about a billion dollars in the U.S. alone<sup>20,21</sup>. Overconfidence in performing this very common surgical procedure and variability in LC operative difficulty have resulted in the scarce implementation of safety guidelines and the consequent non-decreasing incidence of BDI.

Thus, the ubiquity and standardization of LCs have made this procedure an attractive benchmark for CV research and development in minimally invasive surgery<sup>22,23</sup>. In addition, the visual nature and importance of BDI have incentivized both academia and industry to develop CV solutions to solve this well-defined clinical need. Finally, the public release of datasets of annotated LC videos has boosted interest and facilitated research in the field<sup>24</sup>.

### Computer vision analysis

At the coarsest level, a surgery can be described by identifying the procedure being performed. For example, automatic recognition of the type of laparoscopic procedure from the first 10 minutes of surgical procedures has proven highly effective<sup>25</sup>. Though such applications may not immediately seem clinically relevant, they could serve to several indirect purposes, such as reducing annotation efforts for more specific tasks<sup>26</sup> or triggering procedure-specific models without human intervention. Once the type of procedure is identified, consensus suggests that surgical procedures can be described both temporally and spatially using a hierarchy of increasingly detailed descriptors or

annotations (Fig. 1)<sup>27</sup>. In practice, this hierarchy inherently indicates a natural progression of increasingly complex tasks to annotate and model.

At the coarsest temporal level, an entire surgical video can be classified into phases, broad stages of surgical procedures, which can be further broken down into more specific steps that are performed to achieve meaningful surgical goals such as exposing specific anatomic structures. In 2016, EndoNet first tackled the task of surgical phase recognition using a convolutional neural network (CNN) to automatically extract visual features, including information on the appearance of surgical instruments, from LC video frames<sup>24</sup>. A more detailed temporal analysis could be used to recognize specific activities in surgical videos. Initial works on the topic have formalized surgical actions as triplets comprising the tool serving as the end effector, the verb describing the activity at stake, and the anatomy being targeted (e.g., “grasper, retract, gallbladder”)<sup>28</sup>.

At the briefest temporal extreme, the contents of a single frame, such as the instruments or anatomical characteristics, may be described. When applicable, these contents can be further localized spatially, either loosely with markings such as bounding boxes drawn around structures of interest or precisely with segmentation masks delineating objects with pixel-level accuracy. For spatial annotations, the degree of detail is defined by both the type of annotation (e.g., bounding box vs. segmentation masks) and the target being annotated (e.g., tools or tool parts). Further, the relationships between different localized objects can also be described, for example, to describe the interaction or relative position between instruments and anatomical structures.

Invariably, the limiting factor for most clinical applications is the availability of well-annotated datasets. Coarser labels, such as classifying or qualitatively describing the content of a video sequence rather than segmenting each frame, are less cumbersome to annotate but may appear to serve less directly relevant clinical applications. Nevertheless, coarse-grained labels could be used for: (1) data curation and navigation to streamline the use of video for VBA; (2) education by explaining the contents of a video to trainees; and (3) documentation of and navigation to specific data points to later annotate more details.

### Surgical applications

Fundamental work on CV for temporal and spatial analysis of endoscopic videos allowing automated surgical workflow and

scene understanding is being translated to clinically applicable scenarios. LC remains the procedure of choice for demonstrating many such scenarios given its ubiquity and well-defined clinical phenomena; thus, we discuss CV-enabled surgical applications for postoperative video analysis and potential real-time intraoperative assistance in LC. It is important to recognize, however, that such applications are also being investigated for other minimally invasive procedures, gastrointestinal endoscopy, and open surgery<sup>23,29</sup>.

**Quality improvement.** Postoperatively, models for procedure and surgical phase recognition could be used to automatically generate structured and segmented databases to assist with quality improvement initiatives. While such databases would represent an invaluable resource for surgical documentation, research, and education per se, the burden associated with the manual analysis of large quantities of videos presents a considerable bottleneck for adoption. Automated video analysis could be used to digest these large collections of surgical videos, retrieve meaningful video sequences, and extract significant information. For example, full-length surgical videos can be analyzed with phase and tool detection models to identify intraoperative events and effectively produce short videos selectively documenting the division of the cystic duct and the cystic artery, the most critical phase of an LC<sup>30,31</sup>. While this fairly simple approach could be applied to a variety of procedures, adaptation to other use cases would still require considerable development. Very recently, cutting-edge methods have enabled overcoming such barriers by allowing video-to-video retrieval, the task of using a video to search for videos with similar events<sup>32,33</sup>. In addition, models for phase recognition can also be used directly to automatically generate standardized surgical reports of LC. When analyzing such reports based on phase predictions, Berlet et al. found that clusters of incorrectly recognized video frames, i.e. model failures, could indicate complications such as bleeding or problems with gallbladder retrieval<sup>34</sup>. Such events could be linked with the electronic health record to gain insights on patient outcomes after surgery.

**Operative complexity analysis.** CV models can be trained to extract more nuanced information from videos such as surrogates of LC operative difficulty. Since LC operative difficulty correlates with gallbladder inflammation, Loukas et al. trained a CNN to classify the degree of gallbladder wall vascularity yielding performance comparable to expert surgeons<sup>35</sup>. Similarly, Ward et al. trained a CNN to classify gallbladder inflammation according to the Parkland grading scale, a 5-tiered system based on anatomical changes. This classification then contributed to predictions of events such as bile leakage from the gallbladder during surgery and provided insights on how increases in inflammation correlate to prolonged operative times<sup>36</sup>.

**Operative assessment and feedback.** CV models for tool detection have been used to assess the technical skills of surgeons. In this regard, Jin et al. showed that automatically inferred information on tool usage patterns, movement range, and economy correlated with performance assessed by surgeons using validated evaluation metrics<sup>37</sup>. More recently, Lavanchy et al. have proposed to transform automatically extracted tool location information into time-series motion features to use as input of a regression model to predict surgical skills, and distinguish good versus poor technical performance<sup>38</sup>. However, these attempts at automatically assessing technical skills have not been based on existing, validated measures of skill; therefore, more research is required to determine whether automated assessments of skill will supplement or replace traditional assessment methods<sup>39</sup>.

**Intraoperative decision support.** We envision the uptake of AI to assist during minimally invasive procedures (Fig. 2). In this setting, real-time predictions from CV models could be used to guide trainees, enhance surgeon performance, and improve communication in the OR. When starting an LC, CV models could automatically assess the appearance of the gallbladder<sup>35,36</sup>, adjust preoperative estimations of operative difficulty<sup>40</sup>, and suggest whether that case is more appropriate for a trainee or an experienced surgeon. Once the gallbladder is exposed, surgical guidelines suggest using anatomical landmarks to help guide safe zones for incision. For example, Tokuyasu et al. developed a model to automatically detect such key landmarks with bounding boxes<sup>41</sup>.

Similarly, deep learning models could be used to provide a color-coded overlay on the surgical video that could ultimately serve as a navigational assistant for surgeons. Madani et al. have utilized annotations of expert surgeons to train GoNoGoNet to identify safe and unsafe areas of dissection<sup>42</sup>. The endpoint of safe dissection of the hepatocystic triangle is to achieve the critical view of safety (CVS), a universally recommended checkpoint to conclusively identify hepatocystic anatomy and prevent the visual perception illusion causing 97% of major BDIs<sup>43,44</sup>. In this regard, Mascagni et al. have developed a two-stage CV model to first segment surgical tools and fine-grained hepatocystic anatomy to then predict whether each of the three CVS criteria has been achieved<sup>45</sup>.

While automated confirmation of the CVS can provide the surgeon with additional assurance of anatomy, other CV tools can ensure that clips are well placed, and no other structures are inadvertently being clipped. To provide such assistance, Aspart et al. recently proposed ClipAssistNet, a neural network trained to detect the tips of a clip applier during LC<sup>46</sup>. If experienced surgeons may find such assistance unnecessary and even trivial, trainees and early career surgeons may benefit from the reassurance that can be provided by real-time decision-support algorithms such as GoNoGoNet, DeepCVS, and ClipAssistNet. Such algorithms could serve as automated versions of surgical coaches that can facilitate and augment decision-making in the OR<sup>39</sup>.

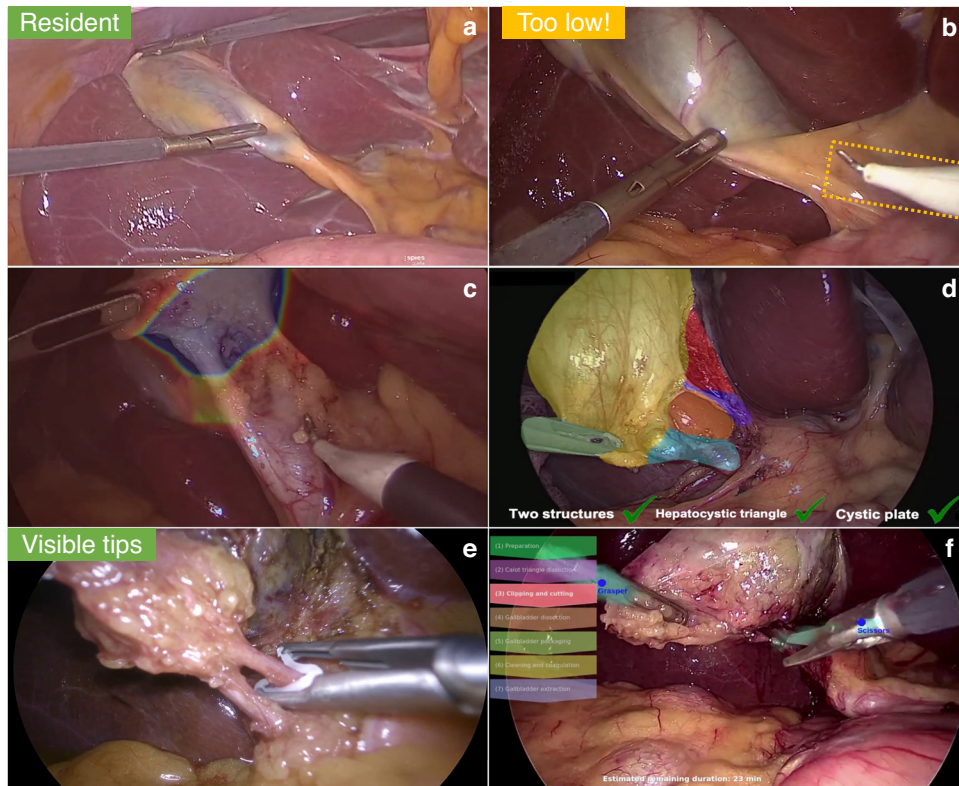
**OR team dynamics.** At a broader level, real-time workflow analysis could be used to improve communication, situational awareness, and readiness of the whole surgical team. Analyzing surgical videos, phase detection models<sup>23</sup> and algorithms to estimate remaining surgical times<sup>47</sup> can help track the progress of the operation to assist OR staff and anesthesia in planning for the current and next case. Furthermore, workflow analysis could help detect deviation from an expected intraoperative course and trigger an automated request for backup or a second opinion. Finally, a visual postoperative summary of the intraoperative events or “surgical fingerprint” could be analyzed with the patient’s preoperative profile to assess the risk of postoperative morbidity or mortality<sup>48</sup>.

## KEY ENABLERS FOR COMPUTER VISION IN SURGERY

Despite the plethora of methods for automated analysis of LC videos presented in the last few years, few AI-based CV systems have been proposed to analyze other surgical procedures, with most focused on minimally invasive procedures. This hinders clinical impact, to the point that no CV application is currently widely used in surgery.

Reasons for this lack of generalization and clinical translation are manifold but largely center around the availability and quality of data and performance of existing modeling approaches, two key elements for CV in surgery which are intimately intertwined.





**Fig. 2 CV-based real-time assistance in laparoscopic cholecystectomy.** Overviewed CV models could be used to evaluate the difficulty of a case and whether it is fit for a surgical resident (a), to warn surgeons against incising below the appropriate site (b), to guide safe dissection (c), to automatically evaluate safety measures (d), to prevent misapplications of clips (e) and to improve OR staff awareness and readiness.

### Surgical data

Historically, surgical procedures were demonstrated in front of trainees and peers in operating theaters with stadium-style seating and windows for natural light. Now, however, operating rooms (ORs) are one of the most siloed components of healthcare systems. Information on OR events is usually only reported in surgeon-dictated post-operative notes or indirectly inferred from postoperative surgical outcomes. As such, it has long been difficult to gather actionable insights on intraoperative adverse events (AE), which occur in up to 2% of all surgical cases<sup>49</sup>. Consequently, clinical needs were mostly identified anecdotally by interviewing surgeons and key opinion leaders, a suboptimal practice prone to biases.

**Variability in surgical data collection.** Today, a greater request for surgical documentation, together with the ease of recording endoscopic videos of minimally invasive surgical procedures, have greatly improved our ability to observe intraoperative events and work toward designing solutions to improve surgical safety and efficiency. However, there is still not much uptake around recording and analyzing surgical data. In a survey of members of a large surgical society, Mazer et al. found surgeons recorded fewer than 40% of their cases though wished up to 80% of videos could be captured. Surgeons felt that lack of equipment, institutional policies, and medico-legal concerns were obstacles to recording cases<sup>50</sup>.

Concerns from surgeons and health systems fearing that intraoperative data might be used against them may be unfounded. A recent review on black box recording devices in the OR has suggested that video data predominantly support surgeons in malpractice cases<sup>51</sup>. Thus, institutions have largely begun to implement an individualized approach to video recording that suits their own needs. Some continue to prohibit

the storage of video, others allow it for select purposes but with specifically outlined parameters (e.g., scheduled destruction of data every 30 days), while others still encourage video recording and storage for quality improvement, education, and research purposes only. Therefore, institutions should engage in a review of existing policies and engage stakeholders such as risk management officers, malpractice insurance carriers, surgeons, and patients to determine the best local strategy for video recording. Clear institutional rules would guide surgeons who wish to record their cases for any number of reasons, including but not limited to use for surgical data science purposes.

**Promoting data acquisition through behavioral incentives.** Policies and incentives may help to further shift the culture of surgical data collection to favor greater operative data collection and use amongst clinicians who may otherwise not consider the value of intraoperative video and computer vision analyses. Institutions that understand the value of video data can play a role in incentivizing clinicians. As an example, AdventHealth, a large academic health system in the United States (US), partnered with a patient safety organization (PSO) to collect and analyze voluntarily submitted data and provides feedback to clinicians, to improve its quality improvement initiatives around operative feedback<sup>52</sup>. In the US, PSOs were established by the Patient Safety and Quality Improvement Act of 2005 and protect the patient safety work products of voluntarily submitted data for quality improvement purposes from civil, criminal, administrative, and disciplinary proceedings except in narrow and specific circumstances. PSOs are organizations that are independent of a health system and certified by the US Agency for Healthcare Research and Quality (AHRQ).

Furthermore, AdventHealth offered continuing medical education (CME) credits necessary for licensing renewals and ongoing

**Table 1.** Common approaches to reduce annotation dependency when learning to perform a task (target task) in a specific domain (target domain).

		Training Annotations		
		Available for target task	Available for different task	Not available
<b>Training data</b>	<b>From target domain</b>	Augmentation	Weakly-supervised learning Multi-task learning	Self-supervised learning Semi-supervised learning
	<b>From different domain</b>	Transfer learning (Domain adaptation)	Transfer learning (e.g. from ImageNet)	

The table reports approaches to facilitate such learning given data from target/different domains, and annotations available for target/different tasks or not available.

board certification as a further individual incentive to surgeons to record and submit videos and review others' videos for quality improvement and educational purposes, such as peer review and feedback. By combining statutory reassurance of privacy with individual incentives in the form of CME, this health system has encouraged voluntary submission of video data from a majority of its surgeons. Such protections and incentives should be considered by other health systems to encourage voluntary participation not just in quality improvement programs but also in efforts to develop CV algorithms that can facilitate such quality improvement initiatives. Ultimately, improved incentives and clearly regulatory guidelines could expand the list of publicly available datasets on which CV algorithms could be developed and tested<sup>53</sup>.

**Limitations in quality of data.** It is not merely the quantity of available data that limits the clinical value of computer vision applications but also the quality of that data. While standardized measurements with predictable variability can be utilized in tabular data, such as laboratory values for hemoglobin or creatinine, defining clinical phenomena in surgical videos (i.e., annotation) can be quite difficult. Open surgery presents unique challenges that occur with occlusion of video data from the surgeon's own movements, necessitating multiple camera angles, additional sensors, or algorithmic approaches to overcome occlusion and consider the added complexity of hand-tool interactions<sup>54–56</sup>.

**Improving data quality.** Clear annotation protocols with extensive annotator training are necessary to ensure that temporal and spatial annotations on surgical videos are clear, reliable, and reproducible. The goals of a given project can help to define the annotation needs and should be clearly established a priori to ensure that appropriate ground truths are established and measured. In addition, annotation protocols should be publicly shared to favor reproducibility and trust by allowing others to collaborate while enabling independent assessment of the ground truth used for training and testing CV models<sup>57</sup>. Ward et al. provide greater detail on the difficulties of annotating surgical video and suggest several key steps that can mitigate against poor or inapplicable model performance related to subpar or inappropriate annotation<sup>58</sup>.

### Artificial intelligence methods

As more and more clinical applications are identified, progressively effective techniques are being introduced to model these applications and bring value to patients. Beyond application-specific modeling, methods are also being developed to help circumvent or mitigate the technical, regulatory, ethical, and clinical constraints endemic to surgery.

**Methods for better leveraging data.** To develop effective clinical solutions, AI models are often trained to replicate expert

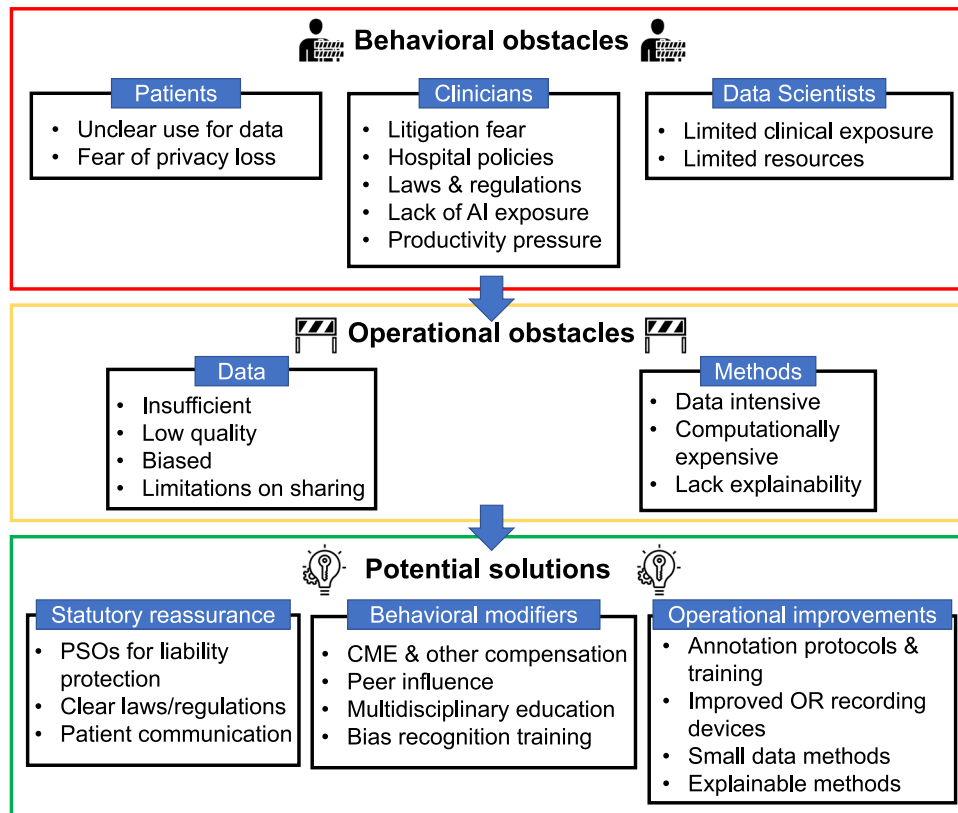
performance from large quantities of well-annotated data (i.e., fully supervised learning). While leading to unprecedented results in medical image analysis<sup>59</sup>, this learning paradigm is highly dependent on the availability of large annotated datasets. Its sustainability is, therefore, severely limited by issues like strict regulatory constraints on data-sharing and the opportunity cost for clinicians to annotate the data, which make the generation of large datasets far from trivial<sup>60</sup>. These issues are further compounded by the need to well-represent and account for variations between patients (anatomy, demographics, etc.), surgeon interactions (workflow, skills, etc.), and OR hardware (instruments, data acquisition systems, etc.).

Several solutions have been explored to increase the amount of data available, such as using synthetically generated datasets<sup>61</sup> or artificially augmenting available annotated datasets<sup>62</sup>. Still, sufficiently modeling the range of possible interactions remains an open problem. Recently, approaches for decentralized training (e.g. federated learning) have begun to gain traction<sup>63</sup>, allowing learning from data at remote physical locations, mitigating privacy concerns, and raising the hope of greater data accessibility.

However, even with large quantities of data available, quality annotations are still scarce and expensive to produce. To reduce the dependency on annotations, different solutions have been proposed, leveraging the intrinsic information present in unlabeled data or repurposing knowledge acquired from different tasks and domains. Self-supervised approaches aim at learning useful information from large amounts of unlabeled data by formulating pre-text tasks which do not require external annotations<sup>64</sup>. Semi-supervised approaches also leverage large quantities of unlabeled data but combine them with small amounts of annotated data. This strategy often involves artificial labeling of unlabeled data, guided by some available labeled data<sup>65,66</sup>.

Weakly supervised methods aim to refine readily available but noisy annotations, such as crowd-sourced labels<sup>67</sup>, or to repurpose existing annotations collected for different tasks (e.g. learning surgical tool localization using non-spatial annotations such as binary tool presence<sup>68</sup>). When such annotations are available concurrently with target-task annotations, multi-task training can be carried out (e.g. using tool presence signals to help inform which surgical phase is being carried out and vice-versa)<sup>24</sup>. Alternatively, transfer-learning approaches help repurpose information learned from different tasks and/or domains, for which annotated datasets are more readily available, and apply it to the domain and task of interest (Table 1). A common example is employing transfer learning from large, well-labeled, non-surgical datasets such as ImageNet<sup>69</sup>. Domain adaptation is another popular transfer-learning paradigm when dealing with data coming from *similar* domains as the target one, such as synthetic surgical datasets<sup>61</sup>.

**Methods for trustworthy AI.** Even as increasingly effective models are being developed for various clinical applications, technical methods are also required to equip surgical staff with the means to explain AI predictions, interpret the reasons behind them,



**Fig. 3 Obstacles and possible solutions for CV in surgery.** Behavior and technical/operational obstacles can limit the development and implementation of CV models in surgery. A combination of statutory, behavioral, and operational changes in the regulatory, clinical, and technical environments could result in improvements in the application of CV for surgery. AI artificial intelligence, PSO patient safety organization, CME continuing medical education, OR operating room.

estimate predictive certainty, and consequently build confidence in the models themselves. These considerations are only now beginning to be addressed in healthcare applications<sup>70</sup> and are particularly glaring in the case of “black-box” algorithms like deep learning-based methods where the relationships between input and output are not always explicit or well-understood. Here, establishing, formalizing, and communicating causal relationships between features of the input and the model output could help mitigate dangerous model failures and potentially inform model design<sup>71</sup>. It is also important to formalize processes to identify, record, and respond to potential sources of error both before and after model deployment. To this end, Liu et al. present a framework for auditing medical artificial intelligence applications<sup>72</sup>.

Future work could look beyond these issues to methods that can identify when dealing with unfamiliar data (out-of-distribution). Aside from enabling clinicians to make informed decisions based on the reliability of the AI system in specific settings, this could also help researchers recognize and address data selection biases and other confounding factors present in the datasets used to train these models.

**Methods for AI translation.** Each clinical application demands specific conditions to be satisfied in order to be delivered in a timely and appropriate manner in line with existing technical and clinical workflows. As several methods are developed to serve and support various stakeholders during different stages of perioperative care, both hardware and software optimizations will also need to be carefully considered. Acceptable latency, errors, and ergonomic interfaces are all key factors in this discussion. For example, certain optimizations such as running these models with reduced precisions may help dramatically reduce the

computational infrastructure needed to deploy these models but may degrade performance. For less time-sensitive applications, cloud computing has been explored for AI-assistance and navigation but is limited by network connectivity<sup>73</sup>.

### ETHICAL, CULTURAL, AND EDUCATIONAL CONSIDERATIONS

The approaches we have reviewed demonstrate that modern methods have the technical capability to translate computer vision advances to surgical care. However, several obstacles and challenges remain to unlock the potential of computer vision in surgery (Fig. 3). While OR translation, clinical validation, and implementation at the scale of CV solutions are surely fundamental to delivering the promised surgical value, these steps involve multiple stakeholders - from device manufacturers to regulators - and remain largely unexplored today. Here we focus on ethical, cultural and educational considerations important to surgeons and their patients.

Several ethical questions must be addressed, including data safety and transparency, privacy, and fairness and bias<sup>74</sup>. Ongoing discussions are occurring at both the national and international levels to determine how best to protect patients without prohibiting innovations in data analysis that could yield safer surgical care. Considerations for data safety, transparency, and privacy include concepts of informed consent by patients, security of data, and data ownership and access, including whether patients have the right to control and oversee how their personal data is being used.



### Patient perspectives on video data

In a qualitative analysis of 49 patient perspectives of video recording via a hypothetical “black box” system that could capture all surgical data in the OR, 88% of patients felt that any ownership of video data belonged to them as opposed to the hospital at which their care was received or to the surgeon who performed their operations<sup>75</sup>. Regulations around ownership, privacy, and use of identifiable and pseudonymized data vary by country (and even by the state, local, and institutional rules) so research efforts have largely been siloed to individual institutions or local consortia where it may be easier to define who owns data under a given legal infrastructure and how it can be used. As efforts continue to better understand the needs of the field in developing technology that could prove lifesaving for surgical care, it will be critically important to ensure that patients are included and prioritized in discussions that concern the use of data generated through their health encounters.

Patients could be a strong advocate for computer vision research in surgery, as many report perceiving that a benefit of video recording is to enable an objective record of the case to assist in future care and serve as medico-legal protection for both the patient *and* the surgeon. Importantly, patients highlighted their desire for such data to be used for continuous quality improvement<sup>75</sup>. The use of computer vision models such as those we have previously described can facilitate each of these benefits today as context-aware algorithms can automatically index cases for rapid review and post hoc use of guidance algorithms can provide visual feedback to surgeons. Indeed, some institutions are using these technologies to facilitate discussions at weekly morbidity and mortality conferences for quality improvement purposes.

### Bias and transparency of datasets

Additional considerations regarding fairness and bias of datasets that affect model performance and lack of algorithmic transparency have also been highlighted in recent publications<sup>76,77</sup>. Bias in datasets must be acknowledged and considered, especially given that many current and future datasets will be obtained from laparoscopic and robotic platforms that may not be as accessible to low- and middle-income countries. It is also important for researchers to recognize that bias can be introduced at the level of each operation, as surgeons carry with them the influence of their training and prior operative experience in surgical decision-making. The amalgamation of such influences will undoubtedly introduce bias into datasets that could impact model performance and thus the generalizability of CV tools in surgery.

### Collaboration to overcome barriers to computer vision research in surgery

As the importance of bias in datasets and the need for representative, generalizable data has been increasingly recognized, efforts have grown around expanding the collaborative nature of AI research for surgery. For example, the Global Surgical Artificial Intelligence Collaborative (GSAC), a nonprofit organization dedicated to promoting the democratization of surgical care through the intersection of education, innovation, and technology, has been facilitating research collaborations across institutions in the US, Canada, and Europe by providing tools for annotation, data sharing, and model development that meets regulatory standards of each of the participating institutions’ home countries. Focused efforts such as GSAC can lower the barrier of entry for institutions and individuals without significant access to either data or computational resources by facilitating cost sharing, providing infrastructure,

and expanding access to both technical and surgical expertise for collaborative work.

### Data science education for clinicians

Finally, education in surgical data science is of paramount importance, both to ensure that current clinicians can understand how computer vision and other AI tools impact their decision-making and patients and to enable future generations to contribute their own insights into developing newer, more sophisticated tools. The Royal College of Physicians and Surgeons of Canada has recently identified digital health literacy as a potential new competency for Canadian physicians in specialty practice, highlighting the importance of new careers that combine medical knowledge with graduate education in AI as well as multidisciplinary clinical teams that incorporate data scientists and AI researchers<sup>78</sup>. A similar conclusion was reached in the UK’s Topol Review on preparing the healthcare workforce for a digital future in the National Health Service (NHS), and the NHS subsequently established Topol Digital Fellowships to teach digital transformation techniques<sup>79</sup>. Institutional, interdisciplinary fellowships are now being established to promote greater clinician literacy in AI topics and greater understanding of clinical problems and workflow by engineers and data scientists. Additionally, institutions such as IHU Strasbourg are offering short, intensive courses in surgical data science to both clinicians and engineers/data scientists to promote interdisciplinary education and collaboration.

### CONCLUSION

Computer vision offers an unprecedented means to study and improve the intraoperative phase of surgery at scale. As both the clinical and data science communities have begun to converge on advancing research and scientific inquiry on how best to utilize CV in surgery, several proof-of-concept applications of potential clinical value have been demonstrated in minimally invasive surgery. Key efforts to generalize such applications focus around streamlining access to surgical data and better modeling methods, always considering the cultural and ethical aspects intrinsic to patient care. As CV in surgery matures, broader societal involvement will be necessary to ensure the promises of CV in surgery are translated safely and efficaciously to assist in the care of surgical patients.

### DATA AVAILABILITY

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Received: 15 July 2022; Accepted: 10 October 2022;  
Published online: 28 October 2022

### REFERENCES

- Weiser, T. G. et al. Estimate of the global volume of surgery in 2012: an assessment supporting improved health outcomes. *Lancet* **385**, S11 (2015).
- Meara, J. G. et al. Global Surgery 2030: Evidence and solutions for achieving health, welfare, and economic development. *Surgery* **158**, 3–6 (2015).
- Childers, C. P. & Maggard-Gibbons, M. Understanding costs of care in the operating room. *JAMA Surg.* **153**, e176233 (2018).
- Zegers, M. et al. The incidence, root-causes, and outcomes of adverse events in surgical units: implication for potential prevention strategies. *Patient Saf. Surg.* **5**, 13 (2011).
- Lewandrowski, K.-U. et al. Regional variations in acceptance, and utilization of minimally invasive spinal surgery techniques among spine surgeons: results of a global survey. *J. Spine Surg.* **6**, S260–S274 (2020).
- Bardakcioglu, O., Khan, A., Aldridge, C. & Chen, J. Growth of laparoscopic colectomy in the United States: analysis of regional and socioeconomic factors over time. *Ann. Surg.* **258**, 270–274 (2013).

7. Richards, M. K. et al. A national review of the frequency of minimally invasive surgery among general surgery residents: assessment of ACGME case logs during 2 decades of general surgery resident training. *JAMA Surg.* **150**, 169–172 (2015).
8. Zhou, M. et al. Effect of haptic feedback in laparoscopic surgery skill acquisition. *Surg. Endosc.* **26**, 1128–1134 (2012).
9. Balvardi, S. et al. The association between video-based assessment of intraoperative technical performance and patient outcomes: a systematic review. *Surg. Endosc.* <https://doi.org/10.1007/s00464-022-09296-6> (2022).
10. Mascagni, P. et al. Intraoperative time-out to promote the implementation of the critical view of safety in laparoscopic cholecystectomy: A video-based assessment of 343 procedures. *J. Am. Coll. Surg.* **233**, 497–505 (2021).
11. Pugh, C. M., Hashimoto, D. A. & Korndorffer, J. R. Jr. The what? How? And Who? Of video based assessment. *Am. J. Surg.* **221**, 13–18 (2021).
12. Feldman, L. S. et al. SAGES Video-Based Assessment (VBA) program: a vision for life-long learning for surgeons. *Surg. Endosc.* **34**, 3285–3288 (2020).
13. Sharma, G. et al. A cadaveric procedural anatomy simulation course improves video-based assessment of operative performance. *J. Surg. Res.* **223**, 64–71 (2018).
14. Ward, T. M. et al. Computer vision in surgery. *Surgery* **169**, 1253–1256 (2021).
15. Hassan, C. et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest. Endosc.* **93**, 77–85.e6 (2021).
16. van Leeuwen, K. G., Schalekamp, S., Rutten, M. J. C. M., van Ginneken, B. & de Rooij, M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. *Eur. Radiol.* **31**, 3797–3804 (2021).
17. Pucher, P. H. et al. Outcome trends and safety measures after 30 years of laparoscopic cholecystectomy: a systematic review and pooled data analysis. *Surg. Endosc.* **32**, 2175–2183 (2018).
18. Törnqvist, B., Strömberg, C., Persson, G. & Nilsson, M. Effect of intended intraoperative cholangiography and early detection of bile duct injury on survival after cholecystectomy: population based cohort study. *BMJ* **345**, e6457 (2012).
19. A prospective analysis of 1518 laparoscopic cholecystectomies. *N. Engl. J. Med.* **324**, 1073–1078 (1991).
20. Rogers, S. O. Jr. et al. Analysis of surgical errors in closed malpractice claims at 4 liability insurers. *Surgery* **140**, 25–33 (2006).
21. Berci, G. et al. Laparoscopic cholecystectomy: first, do no harm; second, take care of bile duct stones. *Surg. Endosc.* **27**, 1051–1054 (2013).
22. Anteby, R. et al. Deep learning visual analysis in laparoscopic surgery: a systematic review and diagnostic test accuracy meta-analysis. *Surg. Endosc.* **35**, 1521–1533 (2021).
23. Garrow, C. R. et al. Machine learning for surgical phase recognition: A systematic review. *Ann. Surg.* **273**, 684–693 (2021).
24. Twinanda, A. P. et al. EndoNet: A deep architecture for recognition tasks on laparoscopic videos. *IEEE Trans. Med. Imaging* **36**, 86–97 (2017).
25. Kannan, S., Yengera, G., Mutter, D., Marescaux, J. & Padoy, N. Future-State Predicting LSTM for early surgery type recognition. *IEEE Trans. Med. Imaging* **39**, 556–566 (2020).
26. Yengera, G., Mutter, D., Marescaux, J. & Padoy, N. Less is more: Surgical phase recognition with less annotations through self-supervised pre-training of CNN-LSTM networks. *arXiv [cs.CV]* (2018).
27. Meireles, O. R. et al. SAGES consensus recommendations on an annotation framework for surgical video. *Surg. Endosc.* In Press, (2021).
28. Nwoye, C. I. et al. Rendezvous: Attention mechanisms for the recognition of surgical action triplets in endoscopic videos. *Med. Image Anal.* **78**, 102433 (2022).
29. Yeung, S. et al. A real-time spatiotemporal AI model analyzes skill in open surgical videos. *Res. Square* <https://doi.org/10.21203/rs.3.rs-1129461/v1> (2021).
30. Mascagni, P. et al. A computer vision platform to automatically locate critical events in surgical videos: Documenting safety in laparoscopic cholecystectomy. *Ann. Surg.* **274**, e93–e95 (2021).
31. Mascagni, P. et al. Multicentric validation of EndoDigest: a computer vision platform for video documentation of the critical view of safety in laparoscopic cholecystectomy. *Surg. Endosc.* <https://doi.org/10.1007/s00464-022-09112-1> (2022).
32. Yu, T. & Padoy, N. Encode the Unseen: Predictive Video Hashing for Scalable Midstream Retrieval. In: Ishikawa, H., Liu, CL., Pajdla, T., Shi, J. (eds) *Computer Vision – ACCV 2020*. ACCV 2020. Lecture Notes in Computer Science, vol 12626. Springer, Cham. [https://doi.org/10.1007/978-3-030-69541-5\\_26](https://doi.org/10.1007/978-3-030-69541-5_26) (2021).
33. Yu, T. et al. Live laparoscopic video retrieval with compressed uncertainty. Preprint at: <https://arxiv.org/abs/2203.04301> (2022).
34. Berlet, M. et al. Surgical reporting for laparoscopic cholecystectomy based on phase annotation by a convolutional neural network (CNN) and the phenomenon of phase flickering: a proof of concept. *Int. J. Comput. Assist. Radiol. Surg.* <https://doi.org/10.1007/s11548-022-02680-6> (2022).
35. Loukas, C., Frountzas, M. & Schizas, D. Patch-based classification of gallbladder wall vascularity from laparoscopic images using deep learning. *Int. J. Comput. Assist. Radiol. Surg.* **16**, 103–113 (2021).
36. Ward, T. M., Hashimoto, D. A., Ban, Y., Rosman, G. & Meireles, O. R. Artificial intelligence prediction of cholecystectomy operative course from automated identification of gallbladder inflammation. *Surg. Endosc.* <https://doi.org/10.1007/s00464-022-09009-z> (2022).
37. Jin, A. et al. Tool Detection and Operative Skill Assessment in Surgical Videos Using Region-Based Convolutional Neural Networks, IEEE Winter Conference on Applications of Computer Vision (WACV), **2018**, pp. 691–699. <https://doi.org/10.1109/WACV.2018.00081> (2018).
38. Lavanchy, J. L. et al. Automation of surgical skill assessment using a three-stage machine learning algorithm. *Sci. Rep.* **11**, 5197 (2021).
39. Ward, T. M. et al. Surgical data science and artificial intelligence for surgical education. *J. Surg. Oncol.* **124**, 221–230 (2021).
40. Vannucci, M. et al. Statistical models to preoperatively predict operative difficulty in laparoscopic cholecystectomy: A systematic review. *Surgery* **171**, 1158–1167 (2022).
41. Tokuyasu, T. et al. Development of an artificial intelligence system using deep learning to indicate anatomical landmarks during laparoscopic cholecystectomy. *Surg. Endosc.* **35**, 1651–1658 (2021).
42. Madani, A. et al. Artificial intelligence for intraoperative guidance. *Ann. Surg.* **276**, 363–369 (2022).
43. Way, L. W. et al. Causes and prevention of laparoscopic bile duct injuries. *Ann. Surg.* **237**, 460–469 (2003).
44. Brunt, L. M. et al. Safe Cholecystectomy Multi-society Practice Guideline and State of the Art Consensus Conference on Prevention of Bile Duct Injury During Cholecystectomy. *Ann. Surg.* **272**, 3–23 (2020).
45. Mascagni, P. et al. Artificial intelligence for surgical safety. *Ann. Surg.* **275**, 955–961 (2022).
46. Aspart, F. et al. ClipAssistNet: bringing real-time safety feedback to operating rooms. *Int. J. Comput. Assist. Radiol. Surg.* **17**, 5–13 (2022).
47. Twinanda, A. P., Yengera, G., Mutter, D., Marescaux, J. & Padoy, N. RSDNet: Learning to predict remaining surgery duration from laparoscopic videos without manual annotations. *IEEE Trans. Med. Imaging* **38**, 1069–1078 (2019).
48. Ward, T. M. et al. Automated operative phase identification in peroral endoscopic myotomy. *Surg. Endosc.* **35**, 4008–4015 (2021).
49. Mavros, M. N. et al. Opening Pandora's box: understanding the nature, patterns, and 30-day outcomes of intraoperative adverse events. *Am. J. Surg.* **208**, 626–631 (2014).
50. Mazer, L., Varban, O., Montgomery, J. R., Awad, M. M. & Schulman, A. Video is better: why aren't we using it? A mixed-methods study of the barriers to routine procedural video recording and case review. *Surg. Endosc.* **36**, 1090–1097 (2022).
51. van Dalen, A. S. H. M., Legemaate, J., Schlack, W. S., Legemate, D. A. & Schijven, M. P. Legal perspectives on black box recording devices in the operating environment. *Br. J. Surg.* **106**, 1433–1441 (2019).
52. United States Code of Federal Regulation. 42 CFR Ch I, Part 3. <https://www.ecfr.gov/current/title-42/chapter-I/subchapter-A/part-3>.
53. Rivas-Blanco, I., Perez-Del-Pulgar, C. J., Garcia-Morales, I. & Munoz, V. F. A review on deep learning in minimally invasive surgery. *IEEE Access* **9**, 48658–48678 (2021).
54. Shimizu, T., Hachiuma, R., Kajita, H., Takatsume, Y. & Saito, H. Hand motion-aware surgical tool localization and classification from an egocentric camera. *J. Imaging* **7**, 15 (2021).
55. Zhang, M. et al. Using computer vision to automate hand detection and tracking of surgeon movements in videos of open surgery. *AMIA Annu. Symp. Proc.* **2020**, 1373–1382 (2020).
56. Goldbraikh, A., D'Angelo, A.-L., Pugh, C. M. & Laufer, S. Video-based fully automatic assessment of open surgery suturing skills. *Int. J. Comput. Assist. Radiol. Surg.* **17**, 437–448 (2022).
57. Mascagni, P. et al. Surgical data science for safe cholecystectomy: a protocol for segmentation of hepatocyst anatomy and assessment of the critical view of safety. Preprint at: <https://arxiv.org/abs/2106.10916> (2021).
58. Ward, T. M. et al. Challenges in surgical video annotation. *Comput Assist Surg. (Abingdon)* **26**, 58–68 (2021).
59. Esteve, A. et al. A guide to deep learning in healthcare. *Nat. Med.* **25**, 24–29 (2019).
60. Maier-Hein, L. et al. Surgical data science - from concepts toward clinical translation. *Med. Image Anal.* **76**, 102306 (2022).
61. Rau, A. et al. Implicit domain adaptation with conditional generative adversarial networks for depth prediction in endoscopy. *Int. J. Comput. Assist. Radiol. Surg.* **14**, 1167–1176 (2019).
62. Shorten, C. & Khoshgoftaar, T. M. A survey on image data augmentation for deep learning. *J. Big Data* **6**, (2019).
63. Kassem, H. et al. Federated cycling (FedCy): Semi-supervised Federated Learning of surgical phases. Preprint at: <https://arxiv.org/abs/2203.07345> (2022).
64. Taleb, A. et al. 3D self-supervised methods for medical imaging. In *Proceedings of the 34th International Conference on Neural Information Processing Systems* (pp. 18158–18172) (2020).



65. Yu, T., Mutter, D., Marescaux, J. & Padoy, N. Learning from a tiny dataset of manual annotations: a teacher/student approach for surgical phase recognition. Preprint at: <https://arxiv.org/abs/1812.00033> (2018).
66. Shi, X., Jin, Y., Dou, Q. & Heng, P.-A. Semi-supervised learning with progressive unlabeled data excavation for label-efficient surgical workflow recognition. *Med. Image Anal.* **73**, 102158 (2021).
67. Zhang, J., Sheng, V. S., Li, T. & Wu, X. Improving crowdsourced label quality using noise correction. *IEEE Trans. Neural Netw. Learn. Syst.* **29**, 1675–1688 (2018).
68. Nwoye, C. I., Mutter, D., Marescaux, J. & Padoy, N. Weakly supervised convolutional LSTM approach for tool tracking in laparoscopic videos. *Int. J. Comput. Assist. Radiol. Surg.* **14**, 1059–1067 (2019).
69. Deng, J. et al. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2009). <https://doi.org/10.1109/cvpr.2009.5206848>.
70. Reyes, M. et al. On the interpretability of artificial intelligence in radiology: Challenges and opportunities. *Radio. Artif. Intell.* **2**, e190043 (2020).
71. Castro, D. C., Walker, I. & Glocker, B. Causality matters in medical imaging. *Nat. Commun.* **11**, 3673 (2020).
72. Liu, X. et al. The medical algorithmic audit. *Lancet Digit Health* **4**, e384–e397 (2022).
73. Sun, L., Jiang, X., Ren, H. & Guo, Y. Edge-cloud computing and artificial intelligence in internet of medical things: Architecture, technology and application. *IEEE Access* **8**, 101079–101092 (2020).
74. Gerke, S., Minssen, T. & Cohen, G. Ethical and legal challenges of artificial intelligence-driven healthcare. In *Artificial Intelligence in Healthcare* 295–336 (Elsevier, 2020).
75. Gallant, J.-N., Brelsford, K., Sharma, S., Grantcharov, T. & Langerman, A. Patient Perceptions of Audio and Video Recording in the Operating Room. *Ann. Surg.* <https://doi.org/10.1097/SLA.0000000000004759> (2021).
76. Gichoya, J. W. et al. AI recognition of patient race in medical imaging: a modelling study. *Lancet Digit Health* **4**, e406–e414 (2022).
77. Pierson, E., Cutler, D. M., Leskovec, J., Mullainathan, S. & Obermeyer, Z. An algorithmic approach to reducing unexplained pain disparities in underserved populations. *Nat. Med.* **27**, 136–140 (2021).
78. Reznick, R. et al. Task Force Report on Artificial Intelligence and Emerging Digital Technologies. Published at: <https://www.royalcollege.ca/rcsite/documents/health-policy/rc-ai-task-force-e.pdf> (2021).
79. The topol review — NHS health education England. *The Topol Review — NHS Health Education England*. Published at: <https://topol.hee.nhs.uk/> (2019).

## ACKNOWLEDGEMENTS

This work was partially supported by French state funds managed by the ANR under references ANR-20-CHIA-0029-01 (National AI Chair AI4ORSafety) and ANR-10-IAHU-02 (IHU Strasbourg). This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813782 - project ATLAS.

## AUTHOR CONTRIBUTIONS

P.M.: Conception and design, drafting and substantial revision. D.A.: Conception and design, drafting and substantial revision. L.S.: Conception and design, drafting and substantial revision. M.S.A.: Design, drafting and substantial revision. A.M.: Design, substantial revision. Y.W.: Design, substantial revision. A.A.: Design, substantial revision. J.R.: Design, substantial revision. S.A.: Design, substantial revision. G.C.: Design, substantial revision. I.B.: Design, substantial revision. N.P.: Conception and design, substantial revision. D.A.H.: Conception and design, drafting and substantial revision. All authors have approved the submitted version and agree to be held personally accountable for the work. P.M. and D.A. contributed equally and share first co-authorship.

## COMPETING INTERESTS

The Authors declare the following Competing Financial Interests: AM is a consultant for Activ Surgical and Genesis MedTech. NP is a scientific advisor for Caresyntax and his laboratory receives a PhD fellowship from Intuitive Surgical. DAH is a consultant for Johnson & Johnson Institute and Activ Surgical. He previously received research support from Olympus Corporation. The Authors declare also the following Competing Non-Financial Interests: PM, MSA, AM, YW, AA, and DAH serve on the board of directors for the Global Surgical AI Collaborative, a non-profit organization that oversees a data sharing and analytics platform for surgical data.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Pietro Mascagni.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022