Author/s:
Badgery, H;Zhou, Y;Siderellis, A;Read, M;Davey, C

Title:
Machine Learning in Laparoscopic Surgery

Date:
2022

Citation:
Badgery, H., Zhou, Y., Siderellis, A., Read, M. & Davey, C. (2022). Machine Learning in Laparoscopic Surgery. Raz, M (Ed.). Nguyen, TC (Ed.). Loh, E (Ed.). Artificial Intelligence in Medicine: Applications, Limitations and Future Directions, (1), pp.175-190. Springer Nature Singapore.

# Machine Learning in Laparoscopic Surgery

**Henry Badgery, Yuning Zhou, Athanasia Siderellis, Matthew Read, and Catherine Davey**

## 1 Introduction

Artificial intelligence (AI) is the use of computers or machines to simulate human behaviour. Machine learning is a subset of AI that involves the use of computer algorithms to learn automatically through experience and exposure to data. Artificial intelligence (AI) technologies are used increasingly in biomedical applications, in part owing to their ability to process and interpret complex images. The field of AI and machine learning in medicine has grown in recent years through access to improved computing power and data collection. In the case of surgery, such advancement has been enabled through the development of improved camera guided minimally invasive and robotic surgical techniques [1]. Artificial neural networks are a

H. Badgery (✉)
Department of HPB/UGI Surgery, St Vincent's Hospital Melbourne,
Melbourne, VIC, Australia
e-mail: henry.badgery@svha.org.au

Y. Zhou · C. Davey
Department of Biomedical Engineering, The University of Melbourne,
Parkville, VIC, Australia
e-mail: yunizhou@student.unimelb.edu.au; catherine.davey@unimelb.edu.au

A. Siderellis
Department of Surgery, The University of Melbourne, St Vincent's Hospital,
Melbourne, VIC, Australia
e-mail: asiderellis@student.unimelb.edu.au

M. Read
Department of HPB/UGI Surgery, St Vincent's Hospital Melbourne,
Melbourne, VIC, Australia

Department of Surgery, The University of Melbourne, St Vincent's Hospital,
Melbourne, VIC, Australia
e-mail: matthew.read@unimelb.edu.au



M. Raz et al. (eds.), *Artificial Intelligence in Medicine*,
https://doi.org/10.1007/978-981-19-1223-8_8

subtype of machine learning algorithm that are modelled on the structure and function of the animal visual cortex. They are able to classify images and identify structures within images based on prior learning and have underpinned a rapid growth in AI technology in surgery [2].

The investigation of machine learning applications in laparoscopic surgery is a fertile area for research. Current applications include the autonomous recognition of anatomy and instruments to provide intraoperative guidance or as a safety mechanism to operating surgeons. In addition, superimposed radiological images, or maps, can even provide an augmented reality environment to help guide the surgeon. Through its ability to track surgical instruments, AI technologies can be used as tools to assess surgical performance and optimise surgical training. Furthermore, it can be used to break down operations into phases as a documentation tool and to improve operating theatre workflow. There is considerable work on exploring ways to apply machine learning to many laparoscopic surgical procedures. Despite the wide potential of AI in laparoscopic surgery, at the time of writing, very few applications have been approved for clinical use.

This chapter explores the use of machine learning algorithms in laparoscopic surgery. It will outline the key machine learning techniques that have been applied to laparoscopy surgery. Preliminary work on machine learning in laparoscopic cholecystectomy will be demonstrated. The ability of neural networks to autonomously recognise anatomical structures and surgical instruments and the clinical relevance of this will be explored. Future applications and hurdles to clinical implementation will also be discussed.

## 2 Machine Learning Applications in Laparoscopic Surgery

Laparoscopic surgery is a minimally invasive operative technique. It involves insufflation of the abdominal cavity with gas to create space followed by the insertion of ports through small incision through the abdominal wall to gain adequate access for both a camera and surgical instruments (Fig. 1). Surgery is then guided by images transmitted to a video monitor (Fig. 2). The use of laparoscopic surgery has increased considerably over the past few decades with the improvement of cameras, video technology, hardware, improved surgical instruments and techniques. It now forms a core competency for surgeons practicing abdominal surgery. The high volume of video data generated through a laparoscopic operation plus the increase in availability of computational power and development of machine learning algorithms has fuelled a rapid increase in the development of machine learning applications in laparoscopic surgery [3].

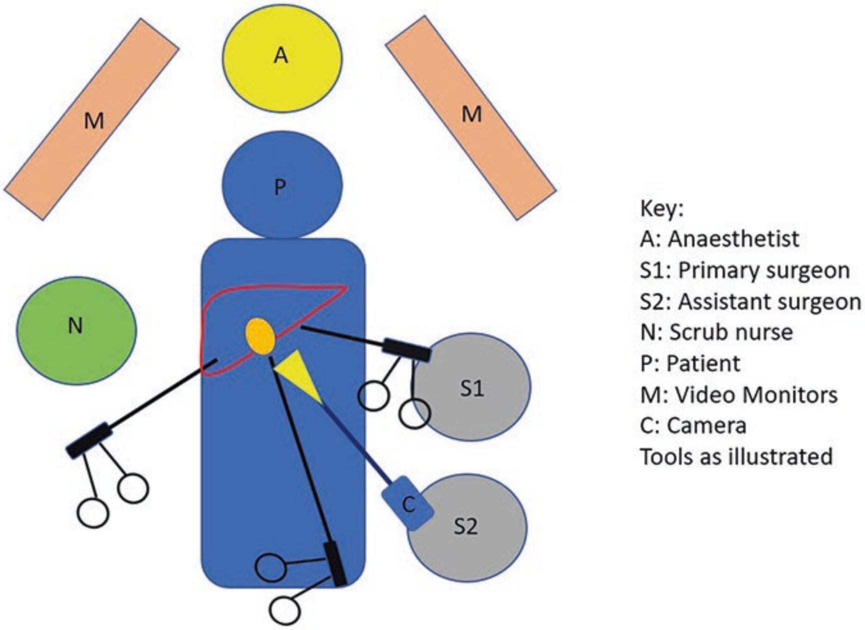


**Fig. 1** Schematic of laparoscopic cholecystectomy set up demonstrating theatre personnel and tool and camera position relative to the liver and gall bladder

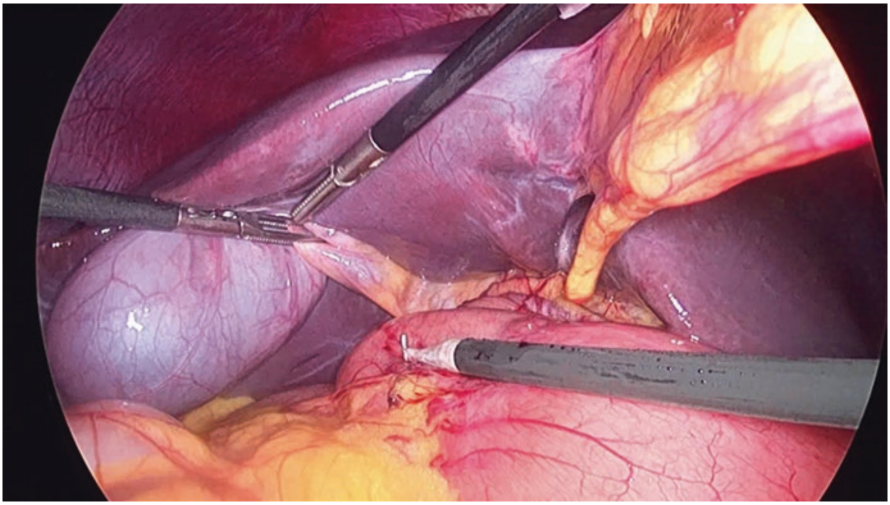

**Fig. 2** Intra-abdominal image from a laparoscopic cholecystectomy

### 2.1 Convolutional Neural Networks

Multiple techniques have been applied to autonomously identify and classify different components of laparoscopic surgery images and videos. The most common and recently utilised algorithms are convolutional neural networks (CNN). CNNs are a type of neural network that aim to mimic the structure and function of the biological visual cortex [2]. These networks consider the hierarchical relationships between neighbouring image pixels rather than only the individual pixel itself. Through learning from previously seen visual image data, they are capable of identifying and classifying objects in different configurations [4]. The volume of relevant visual data that the algorithm is trained on is proportional to its success. CNNs are often combined with other algorithms, which take alternate data inputs and aim to enhance the performance of CNNs with parallel and complementary data when available.

CNNs must be trained on a training dataset. Training datasets are extensive sets of images that are labelled according to the presence of classes, or key features of the image. In the case of laparoscopic surgery images, classes might include anatomical structures, surgical instruments or the presence of absence of disease. Training datasets should include images that the CNN is likely to encounter and interpret and are tailored to the intended output or application of the network. The accuracy of the labels in the training dataset is important as incorrect labels will be perpetuated through the output of the network, reducing the accuracy. Broadly speaking, datasets can be labelled in three ways, classification, detection and segmentation (Fig. 3). Classification requires the binary detection of the presence of absence of a class in an image. Detection refers to locating a class within an image as defined by a bounding box. Segmentation involves the attribution of each pixel to a certain class creating an overlying segmentation mask. Fully supervised labelled data refers to a dataset that has been manually labelled by a human. This is a meticulous and labour-intensive process. Weakly supervised labelled data by contrast involves data that is not labelled explicitly by a human but rather used automated or computer generated labels. These labels are less accurate but easier to obtain. Weakly supervised models are trained with only partially labelled data, allowing the algorithms to extract other features which are considered in classification of inputs.

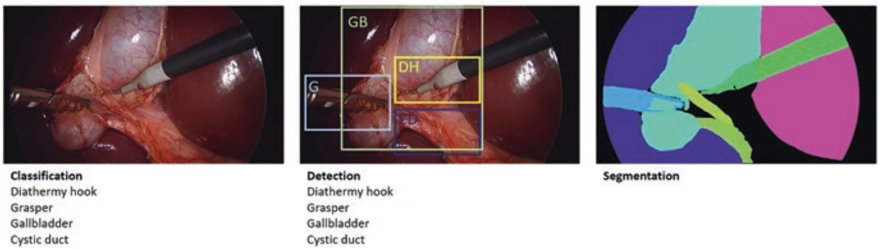


**Fig. 3** Schematic outlining classification, detection with bounding boxes and pixel-wise segmentation labelling

The interpretation and accurate labelling of image data from laparoscopic surgery is complex. The same operation can appear differently between patients and surgeons with variability in pathological states, anatomy and operative technique or approach. The surgical field is often partially occluded by diathermy smoke, blood or fluid. Furthermore, the camera used is non-static. This leads to changes in perspective, angle and zoom plus transient blur or poorly focussed images. While CNNs have successfully been developed capable of identifying phases within an operation, identification of surgical tools and recognition of anatomy, to date, implementation into clinical practice has been limited.

## 3 Surgical Tool Recognition

A significant proportion of the work on machine learning in laparoscopy focusses on surgical tool or instrument detection and segmentation. The detection and tracking of tools is a prerequisite for many clinical applications. These include the assessment of technical performance, understanding surgical phases and workflow, intraoperative decision support as well as warning against potentially unsafe actions. Surgical tool identification can be approached through three ways: binary tool classification, whereby the presence or absence of a tool in an image is recognised; tool detection whereby the presence of the tool is detected and located using a bounding box within the visual field; and segmentation, whereby the precise pixels demonstrating the tool are correctly attributed to that tool, as demonstrated in Fig. 3.

### *3.1 Surgical Tool Classification*

Studies on binary tool classification aim to label images from laparoscopic surgery based on tool presence or absence. Earlier studies have used a superpixel method grouping image pixels together based on similar surrounding information. This allowed the classifier to consider neighbouring information, which improved segmentation compared to a pixel-method [5]. However, there was inherent error in the generation of superpixels which contributed to reduced accuracy of identification. This work contrasts with more recent studies that have employed CNNs with greater success.

Madad Zadeh et al. employed a CNN for binary tool detection during gynaecological surgery with moderate success. As they were limited by a small data set containing a low number of surgical tool annotations, they were only able to achieve an accuracy of 54.5% [6]. In comparison, Kletz et al., using an existing CNN and a much larger dataset, demonstrated superior segmentation and precision yielding a maximum accuracy of 98% [7]. The other binary segmentation studies all used variations of similar CNNs. Lee et al. and Shvets et al. both made comparisons between CNN architectures consisting of different numbers of layers [8, 9]. While these studies tested different modifications of these algorithms, the CNNs with the greatest depth had the best performance in each respective study. They both produced similar analytical metrics for binary segmentation, meaning they both had similar classification capabilities when comparing their outputs to labelled data.

Classification of laparoscopic tools into multiple categories has also been achieved with some success. A number of studies trained and tested CNNs on laparoscopic cholecystectomy videos to classify the same seven or eight instruments into their respective classes [9–16]. Of the studies who reported it [10, 12–14, 16, 17], all achieved relatively poor classification of scissors, with sucker irrigator, bipolar and clipper also often poorly classified. These tools were usually fairly represented in the training datasets. Poor recognition was likely due to their similarity in appearance with other instruments. Images of these tools are demonstrated in Fig. 4.

There is wide variation in the overall accuracy achieved in multi-class tool type detection. Nwoye et al. and Twinanda et al. demonstrated good results with an all-tool presence detection mean average precision of 92.9% and 81% respectively. This was achieved using CNNs incorporating other additional machine learning algorithms [12, 13]. Namazi et al. developed a novel system called LapToolNet using superclasses of instrument combinations and eliminated uncommon instrument combinations. It was trained and validated using existing laparoscopic cholecystectomy video databases (Cholec80 [13] and M2CAI [18]) and outperformed previously published work trained on the M2CAI databases [13, 18, 19]. Sahu et al. by comparison showed poor performance despite utilising additional algorithms with an all tool mean average precision of 54.5% [14]. This is most likely due to the use of a shallower CNN architecture.

Larger datasets are required in more complex procedures or to train algorithms to detect larger sets of tools. Fuentes-Hurtado et al. developed weakly labelled annotated datasets to improve labelling efficiency, applying their method to sleeve gastrectomy and gastric bypass surgeries [15]. The accuracy was 90% of the accuracy of a fully supervised approach suggesting weakly supervised data labelling is viable. CNNs trained on fully supervised segmentation datasets for sleeve gastrectomy and gastric bypass achieved a mean intersection-over-union (mIOU) of 75.4% and 53% respectively [15]. mIOU is a common metric used in machine learning that refers to the common overlap between the ground truth segmentation mask and the CNN generated segmentation mask (Fig. 5).


**Fig. 4** Commonly used laparoscopic instruments

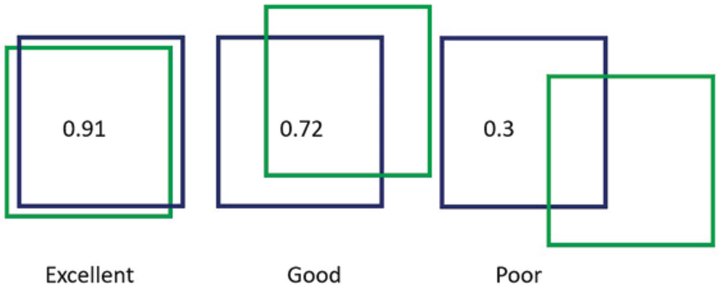


**Fig. 5** Intersection over union. Navy blue = ground truth; green = prediction

### *3.2 Tool Detection and Tracking*

Tool localisation or tracking was often conducted in conjunction with other applications of CNNs, building on previous non-CNN methods. Lee et al. demonstrated their approach had increased tool tracking capabilities compared to instrument tracking using radiofrequency ID tags designed to detect and track instrument position. Radiofrequency ID tags had been the most successful method prior to the application of machine learning techniques [8]. This is thought to be due to reduced interference and more specific tracking due to the classification of tools inherent in these algorithms. Nwoye et al. added long short-term memory network (LSTM) to their CNN which was shown to improve performance compared to the CNN alone [12]. LSTMs bridge time intervals and lags allowing them to enhance the CNNs ability to continuously track the tools in laparoscopic video even as they move in and out of frame [20]. Studies which utilised weakly supervised models had greater tool tracking performance compared to fully supervised models [8, 12, 16]. This improved performance may be through the looser inputs allowing the algorithms to extract additional data or due to these studies having a greater focus on tool tracking. More recently, CNNs have been developed capable of estimating position in a three-dimensional space, known as 3D pose estimation, rather than segmentation on a two dimensional image [21].

The ability of machine learning algorithms to identify and track tools continues to improve through use of problem specific algorithms and larger datasets. It forms a key competency for the implementation of AI technology in a clinical surgical setting.

## 4 Anatomical Structure Recognition

Anatomical structure identification is an important application of machine learning in surgery. This presents additional difficulties over surgical tool detection due to greater variation in appearance between individuals and pathological states. Anatomical structure recognition can be broken down into the binary classification of the presence of organs in a frame, organ detection or recognition of the position within a frame and organ segmentation or the pixel-wise outline of a structure within a frame as well as the automated interpretation or recognition of pathology [3].

Detection of arteries and haemorrhage has been attempted with the aim of prevention or early detection of vessel injury during surgery. Unsupervised algorithms have been applied to this task using recognition of pulse frequency and colour [22–24]. Garcia-Martinez et al. employed a computer vision algorithm to detect haemorrhage in an experimental setting achieving high specificity in the context of low sensitivity [24]. Adams et al. have focused on detection of both visible and obscured arteries for visual enhancement and greater surgical detection through coupling their algorithms to a pulse oximeter, narrowing the frequency range of pulse detection and thereby removing false positives [22]. Akbari et al. coupled pulse oximetry with visual data and was able to detect exposed arteries but was unable to outline margins or categorie vessels [23]. More recent work has explored the assessment of gall bladder wall vascularity using CNN as a surrogate for inflammation and surgical difficulty [25].

The classification of organs has been approached by several methods. Performance was often greater in studies where fewer structures were to be detected. Both Prokopetc et al. and Madad Zadeh et al. demonstrated high detection rates of the uterus (95% and 97% respectively) [6, 26]. Madad Zadeh et al. demonstrated a poor ovarian detection rate (29.6%) likely stemming from greater variation in location and appearance as well as low representation in the training dataset [6]. Both Petscharnig et al. and Moccia et al. attempted classification of multiple structures with limited success [27, 28]. Moccia et al. achieved good results only with addition of a confidence measure, and while they outperformed Petscharnig et al., their algorithm was only applied to porcine models of laparoscopic surgery [27].

Segmentation of anatomy in laparoscopic images has also been explored. Hattab et al. achieved accurate detection of the kidney boundary except in instances of motion blur, obstruction by instruments or otherwise and by the presence of smoke [29]. Similarly, Prokoptec et al. have achieved good detection of the fallopian tube/uterine junction in the context of uterus detection [26]. Greater specificity is achieved probably due to aiding localisation of the area of these landmarks.

Tokuyasu et al. employed an existing CNN to detect four important structures during laparoscopic cholecystectomy, namely the common bile duct, the cystic duct, a landmark of the liver called Rouviere's sulcus and the lower border of the left liver just adjacent to the gallbladder [30]. While average prevision (AP) was relatively low (common bile duct: AP 0.320, cystic duct: 0.074, lower edge of the left medial liver segment: AP 0.314, and Rouviere's sulcus: AP 0.101), subjective

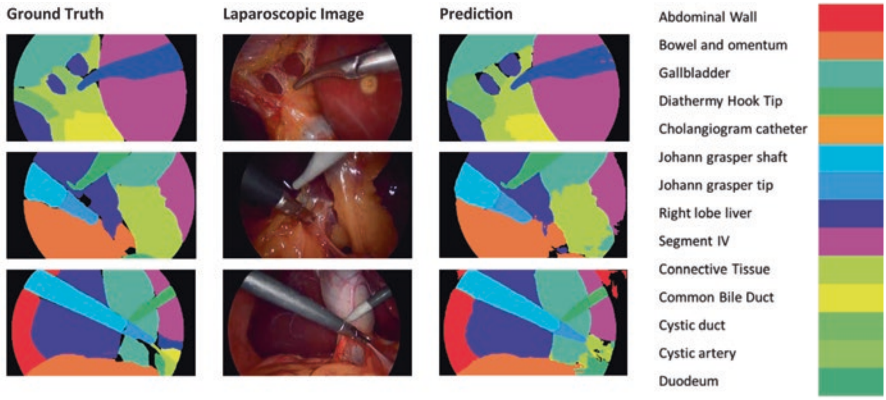


**Fig. 6** Examples of computer-generated segmentation of anatomy and instruments from a laparoscopic cholecystectomy using DeepLabV3+ CNN

validation conducted by surgeons deemed that in 22 out of 23 videos had satisfactory landmark identification. This CNN was trained and tested on non-complicated laparoscopic cholecystectomy videos [30]. Only one study focused on classification of structures within a specific context, namely tagging of different liver views [31].

Accurate detection and segmentation of anatomy is important in the development of AI technology that can help guide surgeons and effectively warn against inadvertent injury. Preliminary work has been successful but is hindered by a paucity of large, labelled datasets. This is compounded by dataset variations in technique, anatomy and pathological states between countries, hospitals, patients and surgeons. International collaborations will improve the performance and generalisability of these networks underpinning the development of effective clinical applications (Fig. 6).

## 5 Surgical Phase Recognition

Classification of laparoscopic surgery videos into pre-determined surgical phases is another important application for machine learning. Accurate recognition of surgical phase can improve the workflow of the operating theatre through improved timing and anticipated staffing and equipment requirements. Anaesthetic use can be minimised through improved prediction of surgical duration. Furthermore, video documentation can be improved contributing to surgical training tools. AI powered surgical phase recognition has been used in clinical products to catalogue and analyse surgical videos for surgical training, research and the creation and maintenance of catalogued surgical videos. This is one of the few areas where AI has been implemented in clinical settings in laparoscopic surgery.

### *5.1 Surgical Phase Recognition in Laparoscopic Cholecystectomy*

A number of studies have achieved surgical phase detection in laparoscopic cholecystectomy, often using pre-existing training datasets [3, 14, 32]. In addition to CNN interpretation of visual data, several studies have employed additional inputs to improve performance. Dergachyova et al. used both visual and instrument use data as inputs, providing additional temporal information which improved the detection of surgical phase [32]. Similarly, Jin et al. improved their algorithm with the addition of a LSTM, outperforming Twinanda et al. on the Cholec80 dataset, and openly available laparoscopic cholecystectomy video dataset [13, 33].

### *5.2 Surgical Phase Recognition in Other Laparoscopic Procedures*

Surgical phase recognition has also been applied to other surgical procedures such as sleeve gastrectomy and colorectal surgery. Each of the studies involving sleeve gastrectomy had very different approaches. Hashimoto et al. and Volkov et al. have applied additional networks to CNNs for this task [34, 35]. Volkov et al. obtained greater performance on their dataset, however the surgical phases all occurred in the same order in each video, which may suggest the algorithm used by Volkov et al. *is not as* variable and generalised. Surgical phase classification in colorectal surgery had more variable results, probably due to the greater number of surgical phases and variability between surgeons and operations and therefore the requirement of larger training datasets. Jalal et al. demonstrated low performance in their use of AlexNet, a CNN to classify sigmoid resection into ten phases [36]. Compared to Jalal et al., Kitaguchi et al. were able to obtain very high overall accuracy for the classification of sigmoid resection into eleven phases [37]. This may be attributed to the use of a light gradient boosting machine, an efficient machine learning algorithm based on decision tree classification which enhances the classification accuracy of the CNN. Additionally, Kitaguchi et al. had large annotated data sets in both of their studies, which also may have contributed to their high classification accuracy [37, 38].

## 6 Machine Learning Applications Specific to Laparoscopic Cholecystectomy

Laparoscopic cholecystectomy has become a commonly used operation in the exploration of machine learning applications due to its common performance and therefore ample data availability plus the relatively standardised nature of the procedure. The use of machine learning in laparoscopic cholecystectomy has the propensity to improve safety through improved anatomical structure identification, the recognition of an unsafe line of dissection, early recognition of bile duct injury (BDI), a major potential complication of cholecystectomy, improved recognition of important steps and landmarks and through training and performance assessment.

Madani et al. designed three CNNs capable of demarcating the 'go-zone', 'no-go zone' as well as achieving limited segmentation of anatomy (liver, hepatocystic triangle and gallbladder) in laparoscopic cholecystectomy surgery [39]. These zones constitute areas where ongoing dissection has a low risk or high risk respectively of inadvertent injury to critical structures. The study was limited by the dataset with the use of 10 randomly selected labelled frames per video. Furthermore, anatomical recognition was crude and unable to identify finer structures and recognise anatomical variation. A heat map was employed based on correlation between independently labelled images by a panel of surgeons. Each surgeon was asked to independently assess their opinion on the 'safe' and 'unsafe' zone. Inevitably, there was a degree of variability between each surgeon's assessment that was presented as probability heat map. This provided a visual probability of each pixel belonging to a zone rather than absolute demarcation of 'ground truth'. This is a highly labour-intensive process however and may not be suitable for the development of large, broader datasets [39].

The work by Tokuyasu et al., as discussed earlier, demonstrates the feasibility of a landmark identification system. The most common cause of BDI is misidentification of biliary anatomy by the surgeon [40]. This is often due to lack of operator awareness and error, anatomical variation between patients, gallbladder traction altering the perception of the anatomy, and inflammation and scarring of the soft tissue causing distortion of the anatomy [41]. This technology has the potential, not only to aid the training surgeon in anatomical recognition, but also may aid the experienced surgeon in interrupting unsafe plan continuation or tunnel vision and confirmation biases underlying potential BDI [30]. Confirmation bias is a cognitive error whereby one preferentially seeks information that supports what one thinks to be the case and ignores information that suggests otherwise.

The use of the critical view of safety (CVS) has been an important development in the safety on laparoscopic cholecystectomy [42]. The CVS is a goal of dissection during a laparoscopic cholecystectomy where all connective tissues is cleared from the area between the liver and gallbladder and the artery and duct to the gall bladder are clearly and unmistakably defined. By achieving the CVS, the surgeon can be confident that the perceived anatomical configuration is correct and that no inadvertent bile duct injury has or will take place. Mascagni et al. developed a computer vision model capable of locating and labelling critical events during cholecystectomy, in particular the CVS and cystic duct division to improve surgical documentation and reinforce the need for surgeons to achieve a satisfactory degree of dissection [43].

Machine learning algorithms can also serve as an adjunct to interpretation of intraoperative imaging, as demonstrated in Fig. 7. Automated interpretation and documentation plus recognition of filling defects or presence of abnormal stones

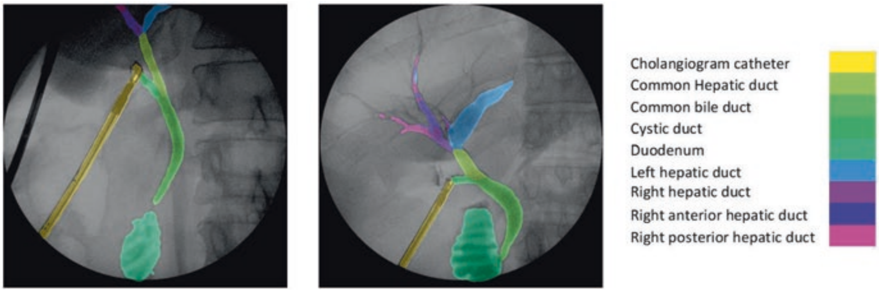


**Fig. 7** Machine learning segmentation of intraoperative cholangiogram performed as part of a CNN. Colour overlay represents the CNN generated segmentation mask

and anatomical aberrancy in intraoperative cholangiograms is another promising application of machine learning in laparoscopic cholecystectomy.

## 7 Challenges in Machine Learning and Laparoscopic Surgery

While the field of machine learning in laparoscopic surgery has progressed considerably, there remain considerable challenges. One major limitation is the size, variability, accessibility and quality of existing datasets for the training, testing and validation of these algorithms. CNNs require large training datasets. The labelling of these datasets is a labour intensive and time-consuming process. As described, some existing studies have attempted to improve labelling efficiency by applying weak labels, then training their algorithms to deduct the information that more extensive labels would have provided [8, 12, 15, 16, 44]. This method however requires larger datasets than fully supervised training and generally yields inferior results in most applications. Many existing studies have used the same set of publicly available datasets, limiting the applicability to other settings [13, 18]. The availability of more high quality, multi-site datasets from a variety of procedures and pathologies will improve the capabilities of machine learning in laparoscopic surgery [2].

Another challenge is in consistency of data labelling and how to define the ground truth. It is often not possible to accurately demarcate adjacent anatomical structures, for example the point at which lower gallbladder becomes the cystic duct. Similarly, the presence of adipose tissue, blood and fascia overlying structures further complicates this. Furthermore, the notion of ground truth is contentious. The source of ground truth in visual data is generally derived from human labelling and subject to an inevitable degree of bias. There will always be degree of interobserver variability between labellers with inevitable inaccuracies within the training datasets. Better defined methodologies and consensus for the labelling of 'ground truth' will improve datasets and in turn performance of ML algorithms. Furthermore, the assessment of success of the validation of machine learning algorithms is not consistent. An arbitrary threshold of 0.5 has been used by a number of studies for correct segmentation. While this is the accepted standard in the machine learning literature more broadly, it is not validated for clinical and surgical applications [3].

The published work on machine learning laparoscopic surgery generally applies to more straightforward operative cases on a limited number of surgical applications. While these provide important prototypes, the clinical value of these CNNs is limited as it is in more complex surgical cases or in more advanced pathology where effective machine learning algorithms are likely to translate to improved clinical outcomes.

Finally, most of the aforementioned machine learning applications in laparoscopic surgery have not been applied clinically. Much of the recent work focusses on methodology from a computational point of view. There is limited reporting on surgical indications, pathological degrees, patient characteristics and surgeon preferences. Furthermore, most studies are retrospective with significant heterogeneity with minimal data demonstrating AI in laparoscopic surgery leading to improved clinical outcomes. A greater focus on clinical application, plus collaboration between clinicians and computer scientists will be important in the further development and translation of this technology.

## 8 Conclusion

The application of machine learning to laparoscopic surgery is a growing and promising field with the potential to revolutionise surgical practice. Applications have included surgical tool detection and classification, anatomical structure classification, landmark classification and surgical phase detection. These have been achieved with varying levels of success. While some studies show high accuracy and precision in their algorithms to perform these tasks, none have applied them to real-life surgical procedures. There also appears to be limited research in utilising these methods in anatomical structure and landmark classification, identifying a need for further study. The field is rapidly progressing with incremental improvements in algorithm performance and novel applications being regularly published in the literature. As datasets are developed, labelling guidelines finetuned and international collaborations are fostered, new clinical applications can be developed and validated underpinning advancement in laparoscopic surgery.

## References


1. Garrow CR, Kowalewski KF, Li L, Wagner M, Schmidt MW, Engelhardt S et al (2020) Machine learning for surgical phase recognition: a systematic review. Ann Surg
2. Hashimoto DA, Rosman G, Rus D, Meireles OR (2018) Artificial intelligence in surgery: promises and perils. Ann Surg 268(1):70–76
3. Anteby R, Horesh N, Soffer S, Zager Y, Barash Y, Amiel I et al (2021) Deep learning visual analysis in laparoscopic surgery: a systematic review and diagnostic test accuracy meta-analysis. Surg Endosc
4. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9(4):611–629
5. Bodenstedt S, Görtler J, Wagner M, Kenngott H, Müller B, Dillmann R et al (2016) Superpixel-based structure classification for laparoscopic surgery, p 978618
6. Madad Zadeh S, Francois T, Calvet L, Chauvet P, Canis M, Bartoli A et al (2020) SurgAI: deep learning for computerized laparoscopic image understanding in gynaecology. Surg Endosc 29:29
7. Kletz S, Schoeffmann K, Husslein H (2019) Learning the representation of instrument images in laparoscopy videos. Healthc Technol Lett 6(6):197–203
8. Lee EJ, Plishker W, Liu X, Bhattacharyya SS, Shekhar R (2019) Weakly supervised segmentation for real-time surgical tool tracking. Healthc Technol Lett 6(6):231–236
9. Shvets AA, Rakhlin A, Kalinin AA, Iglovikov VI Automatic instrument segmentation in robot-assisted surgery using deep learning. 2018 17th IEEE international conference on machine learning and applications (ICMLA), 17–20 December, 2018. 2018
10. Jin A, Yeung S, Jopling J, Krause J, Azagury D, Milstein A et al (2018) Tool detection and operative skill assessment in surgical videos using region-based convolutional neural networks, pp 691–699
11. Kamrul Hasan SM, Linte CA (2019) U-NetPlus: a modified encoder-decoder U-net architecture for semantic and instance segmentation of surgical instruments from laparoscopic images. In: Conference proceedings: annual international conference of the IEEE Engineering in Medicine & Biology Society, pp 7205–7211
12. Nwoye CI, Mutter D, Marescaux J, Padoy N (2019) Weakly supervised convolutional LSTM approach for tool tracking in laparoscopic videos. Int J Comput Assist Radiol Surg 14(6):1059–1067
13. Twinanda AP, Shehata S, Mutter D, Marescaux J, de Mathelin M, Padoy N (2017) EndoNet: a deep architecture for recognition tasks on laparoscopic videos. IEEE Trans Med Imaging 36(1):86–97
14. Sahu M, Mukhopadhyay A, Szengel A, Zachow S (2016) Tool and phase recognition using contextual CNN features. ArXiv:abs/1610.08854
15. Fuentes-Hurtado F, Kadkhodamohammadi A, Flouty E, Barbarisi S, Luengo I, Stoyanov D (2019) EasyLabels: weak labels for scene segmentation in laparoscopic videos. Int J Comput Assist Radiol Surg 14(7):1247–1257
16. Vardazaryan A, Mutter D, Marescaux J, Padoy N (eds) (2018) Weakly-supervised learning for tool localization in laparoscopic videos. Intravascular imaging and computer assisted stenting and large-scale annotation of biomedical data and expert label synthesis. Springer, Cham
17. Bareum C, Kyungmin J, Songe C, Jaesoon C (2017) Surgical-tools detection based on convolutional neural network in laparoscopic robot-assisted surgery. In: Conference proceedings: annual international conference of the IEEE Engineering in Medicine & Biology Society. 2017, pp 1756–1759
18. Stauder R, Ostler D, Kranzfelder M, Koller S, Feußner H, Navab N (2016) The TUM LapChole dataset for the M2CAI 2016 workflow challenge. ArXiv:abs/1610.09278
19. Namazi B, Sankaranarayanan G, Devarajan V (2021) A contextual detector of surgical tools in laparoscopic videos using deep learning. Surg Endosc
20. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780

21. Hasan MK, Calvet L, Rabbani N, Bartoli A (2021) Detection, segmentation, and 3D pose estimation of surgical tools using convolutional neural networks and algebraic geometry. Med Image Anal 70:101994
22. Adams F, Schoelly R, Schlager D, Schoenthaler M, Schoeb DS, Wilhelm K et al (2017) Algorithm-based motion magnification for video processing in urological laparoscopy. J Endourol 31(6):583–587
23. Akbari H, Kosugi Y, Khorgami Z (2009) Image-guided preparation of the Calot's triangle in laparoscopic cholecystectomy. In: Conference proceedings: annual international conference of the IEEE Engineering in Medicine & Biology Society, pp 5649–5652
24. Garcia-Martinez A, Vicente-Samper JM, Sabater-Navarro JM (2017) Automatic detection of surgical haemorrhage using computer vision. Artif Intell Med 78:55–60
25. Loukas C, Frountzas M, Schizas D (2021) Patch-based classification of gallbladder wall vascularity from laparoscopic images using deep learning. Int J Comput Assist Radiol Surg 16(1):103–113
26. Prokopetc K, Collins T, Bartoli A (2015) Automatic detection of the uterus and fallopian tube junctions in laparoscopic images. Inf Process Med Imag 24:552–563
27. Moccia S, Wirkert SJ, Kenngott H, Vemuri AS, Apitz M, Mayer B et al (2018) Uncertainty-aware organ classification for surgical data science applications in laparoscopy. IEEE Trans Biomed Eng 65(11):2649–2659
28. Petscharnig S, Schöffmann K (2018) Learning laparoscopic video shot classification for gynecological surgery. Multimed Tools Appl 77(7):8061–8079
29. Hattab G, Arnold M, Strenger L, Allan M, Arsentjeva D, Gold O et al (2020) Kidney edge detection in laparoscopic image data for computer-assisted surgery: kidney edge detection. Int J Comput Assist Radiol Surg 15(3):379–387
30. Tokuyasu T, Iwashita Y, Matsunobu Y, Kamiyama T, Ishikake M, Sakaguchi S et al (2020) Development of an artificial intelligence system using deep learning to indicate anatomical landmarks during laparoscopic cholecystectomy. Surg Endosc 18:18
31. Nazir A, Cheema MN, Sheng B, Li P, Li H, Yang P et al (2020) SPST-CNN: spatial pyramid based searching and tagging of liver's intraoperative live views via CNN for minimal invasive surgery. J Biomed Inform 106:103430
32. Dergachyova O, Bouget D, Huaulme A, Morandi X, Jannin P (2016) Automatic data-driven real-time segmentation and recognition of surgical workflow. Int J Comput Assist Radiol Surg 11(6):1081–1089
33. Jin Y, Dou Q, Chen H, Yu L, Qin J, Fu C-W et al (2017) SV-RCNet: workflow recognition from surgical videos using recurrent convolutional network. IEEE Trans Med Imaging:1
34. Hashimoto DA, Rosman G, Witkowski ER, Stafford C, Navarette-Welton AJ, Rattner DW et al (2019) Computer vision analysis of intraoperative video: automated recognition of operative steps in laparoscopic sleeve gastrectomy. Ann Surg 270(3):414–421
35. Volkov M, Hashimoto D, Rosman G, Meireles O, Rus D (2017) Machine learning and coresets for automated real-time video segmentation of laparoscopic and robot-assisted surgery
36. Jalal N, Alshirbaji T, Möller K (2018) Evaluating convolutional neural network and hidden Markov model for recognising surgical phases in sigmoid resection. Curr Dir Biomed Eng 4:415–418
37. Kitaguchi D, Takeshita N, Matsuzaki H, Takano H, Owada Y, Enomoto T et al (2019) Real-time automatic surgical phase recognition in laparoscopic sigmoidectomy using the convolutional neural network-based deep learning approach. Surg Endosc 03:03
38. Kitaguchi D, Takeshita N, Matsuzaki H, Oda T, Watanabe M, Mori K et al (2020) Automated laparoscopic colorectal surgery workflow recognition using artificial intelligence: experimental research. Int J Surg 79:88–94
39. Madani A, Namazi B, Altieri MS, Hashimoto DA, Rivera AM, Pucher PH et al (2020) Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. Ann Surg

40. Connor S, Garden OJ (2006) Bile duct injury in the era of laparoscopic cholecystectomy. BJS 93(2):158–168
41. Jabłońska B, Lampe P (2009) Iatrogenic bile duct injuries: etiology, diagnosis and management. World J Gastroenterol 15(33):4097–4104
42. Strasberg SM, Brunt LM (2010) Rationale and use of the critical view of safety in laparoscopic cholecystectomy. J Am Coll Surg 211(1):132–138
43. Mascagni P, Alapatt D, Urade T, Vardazaryan A, Mutter D, Marescaux J et al (2021) A computer vision platform to automatically locate critical events in surgical videos: documenting safety in laparoscopic cholecystectomy. Ann Surg 274:e93–e95
44. Yengera G, Mutter D, Marescaux J, Padoy N (2018) Less is more: surgical phase recognition with less annotations through self-supervised pre-training of CNN-LSTM networks