

Artificial Intelligence for Real-Time Verification of the Critical View of Safety in Laparoscopic Cholecystectomy

James O'Leary^{1*}, Aisling Dunne¹, Sean O'Brien²

¹ Department of Hepatopancreatobiliary Surgery and AI-Assisted Surgery, School of Medicine, Trinity College Dublin, Dublin, Ireland.

² Department of Surgical Safety and Intraoperative Decision Support, School of Medicine, University College Cork, Cork, Ireland.

*E-mail ✉ oleary@tcd.ie

Abstract

The Critical View of Safety (CVS) is the established method to minimize bile duct injury risk during Laparoscopic Cholecystectomy (LC). This investigation assessed the precision and real-world functionality of a pre-trained AI model for confirming successful CVS attainment in elective LC surgeries within an actual operating theatre environment. A deep learning neural network, earlier trained on annotated footage from 700 LC procedures to recognize CVS elements, was connected directly to the laparoscopic imaging system in the operating room. This allowed ongoing, instantaneous assessment of CVS during elective cases. AI-generated feedback was systematically compared against the surgeon's intraoperative evaluation, alongside documentation of overall CVS success rates. Across 40 successive LC operations, surgeons indicated CVS achievement in 34 cases (85%). In the 6 procedures where CVS could not be obtained owing to intense inflammation or distorted anatomy, the AI system aligned completely with the surgeons' judgment and correctly refrained from confirming CVS. Of the 34 cases reported as successful by surgeons, the AI model correctly detected CVS in 33 instances. This yielded a specificity of 100% [95%-CI 98.1%, 100%] and sensitivity of 97% [95%-CI 96.1%, 98.2%] for CVS detection. The trained AI model accurately identified CVS in elective LC with exceptional reliability in real-time operating room scenarios. Its deployment may promote greater consistency in CVS attainment, leading to safer procedures and better patient outcomes.

Keywords: Artificial intelligence system, AI in surgery, Laparoscopic cholecystectomy, Machine learning

Introduction

Algorithms based on computer vision, Artificial Intelligence (AI), and Machine Learning (ML) have achieved notable success in diverse medical fields. These include analysis of radiological imaging, pathological diagnosis from tissue samples, and polyp detection in colonoscopies [1-4]. However, their application in surgery for real-time intraoperative assistance and safety enhancement remains an active area of research, with initial positive findings [1, 4-7].

As one of the most routine laparoscopic operations, Laparoscopic cholecystectomy (LC) has drawn increasing attention for the use of AI to recognize important procedural steps from video recordings [4, 5, 8].

Our earlier research established that an AI system can determine LC phases with high precision, including in difficult cases, relative to expert manual labeling [9]. Despite a bile duct injury (BDI) occurrence rate in LC of only 0.3–1.5% [10-15], this figure exceeds the rates from the open surgery period and has shown little decline over many years. This persistence continues despite greater expertise, better equipment, and technical advancements [10-15].

BDI most often results from errors in visual interpretation of anatomy, particularly in regions known for structural variability [16].

Access this article online

<https://smerpub.com/>

Received: 18 August 2025; Accepted: 11 December 2025

Copyright CC BY-NC-SA 4.0

How to cite this article: O'Leary J, Dunne A, O'Brien S. Artificial Intelligence for Real-Time Verification of the Critical View of Safety in Laparoscopic Cholecystectomy. J Med Sci Interdiscip Res. 2025;5(2):143-50. <https://doi.org/10.51847/zyGCA6EY14>

To reduce BDI incidence, the Critical View of Safety (CVS) technique—mandating its clear achievement before transecting the cystic structures—was recommended [17-20].

In recent years, multiple AI-ML frameworks have emerged for CVS recognition in routine and demanding LC scenarios. One method, DeepCVS, processes images and segmentation masks through a compact neural network to predict CVS elements [9]. Alternative work has emphasized binary segmentation of the cystic artery and duct [21-27]. Other efforts combined Convolutional Neural Networks (CNNs) with Recurrent Neural Networks (RNNs) to verify the standard three CVS criteria [28], while another project treated CVS detection as an attention-guided Multi-Instance Learning challenge [29].

Building upon these advancements, our method follows a similar foundation to DeepCVS but incorporates temporal aggregation and supports real-time functionality.

This study examines the effectiveness and precision of the developed AI model for confirming CVS attainment during LC in a live operating room. Findings from this and future work may support integrating such technology to guide surgical decisions and ultimately enhance patient safety and outcomes.

Materials and Methods

Training/evaluation dataset

The dataset consisted of 700 cholecystectomy recordings: 620 collected from five Israeli hospitals and 80 sourced from the open-access Cholec80 collection [8]. Institutional videos spanned November 1, 2010, to October 1, 2020. Procedures were included for patients 18 years or older undergoing LC due to biliary colic, acute, or chronic cholecystitis. We used an 80:20 split for training and testing, dividing by entire case rather than individual frames to prevent overlap.

Mirroring an earlier approach [27], we isolated 60-second clips ending just before cystic structure division. Three senior surgeons reviewed each segment for CVS compliance according to the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) standards, evaluating these four criteria:

- C1: Structure criterion: Only the two pertinent tubular structures entering the gallbladder are visible.

- C2: Hepatocystic Triangle: The hepatocystic triangle is largely cleared of fat and connective tissue.
- C3a: Cystic plate criterion: The inferior aspect of the gallbladder is freed from the cystic plate.
- C3b: Liver criterion: Liver tissue appears between the two tubular structures.

Splitting C3 into sub-criteria C3a and C3b helped raise agreement levels among evaluators.

A Deep Learning neural network was then trained on the labeled segments to detect relevant hepatocystic anatomy and classify CVS fulfillment.

Deep learning model architecture

The model operates through three distinct phases. For every endoscopic frame, the initial phase segments hepatocystic structures. The middle phase evaluates the likelihood of each CVS criterion using the segmented data. The final phase combines these assessments temporally to output the overall CVS status.

First stage: hepatocystic anatomy segmentation

Targeted structures included the cystic duct, cystic artery, intervening liver tissue, and cystic plate. Segmentation relied on a SpineNet-49 backbone [21]. Pretraining occurred on the COCO natural image dataset [30], followed by fine-tuning of the full model on 2,500 surgeon-annotated images created internally by three specialists.

Second stage: CVS criteria estimation

A ResNet50-based classification network [22] receives the raw image paired with the first-stage segmentation output and generates probability scores for each of the four CVS criteria.

Third stage: CVS criteria aggregation

This stage smooths criterion probabilities across time with a median filter over a 0.5-second window. Averaged filtered values then determine CVS status via dual hysteresis thresholds (0.6 for shifting to 'CVS' and 0.4 for shifting to 'No CVS'). The hysteresis design minimizes instability and rapid oscillations.

Training and validation followed the previously outlined dataset. Frame processing time averaged 55 ms, enabling real-time analysis by evaluating alternate frames in a 30 FPS stream.

Study

With institutional IRB clearance obtained, the CVS detection platform was tested prospectively in real operating room conditions.

Hardware was interfaced with the standard laparoscopic camera output across 40 elective LC cases, with the AI-ML software operating silently in the background (Figure 1). CVS achievement required meeting at least 3

of 4 criteria. Operating room staff recorded the surgeon-declared CVS moment. Although processed live, AI predictions remained hidden from the surgical team. Postoperative analysis matched surgeon timestamps against system outputs. Figure 2 summarizes the study process.

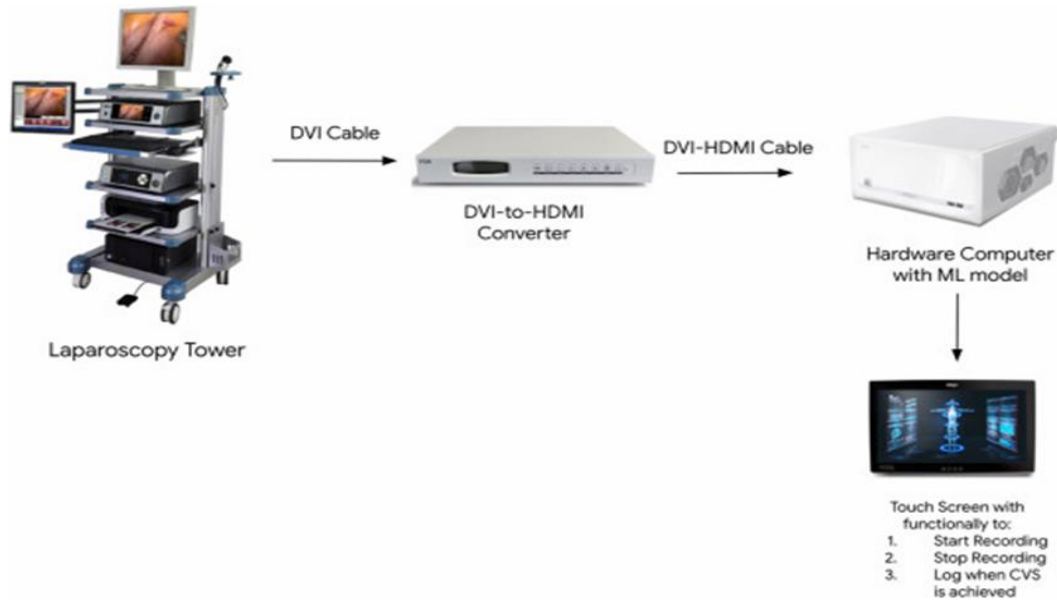


Figure 1. Overview of the system setup. A dedicated hardware unit links directly to the regular laparoscopic camera video feed, while background software executes the AI-ML model for processing.

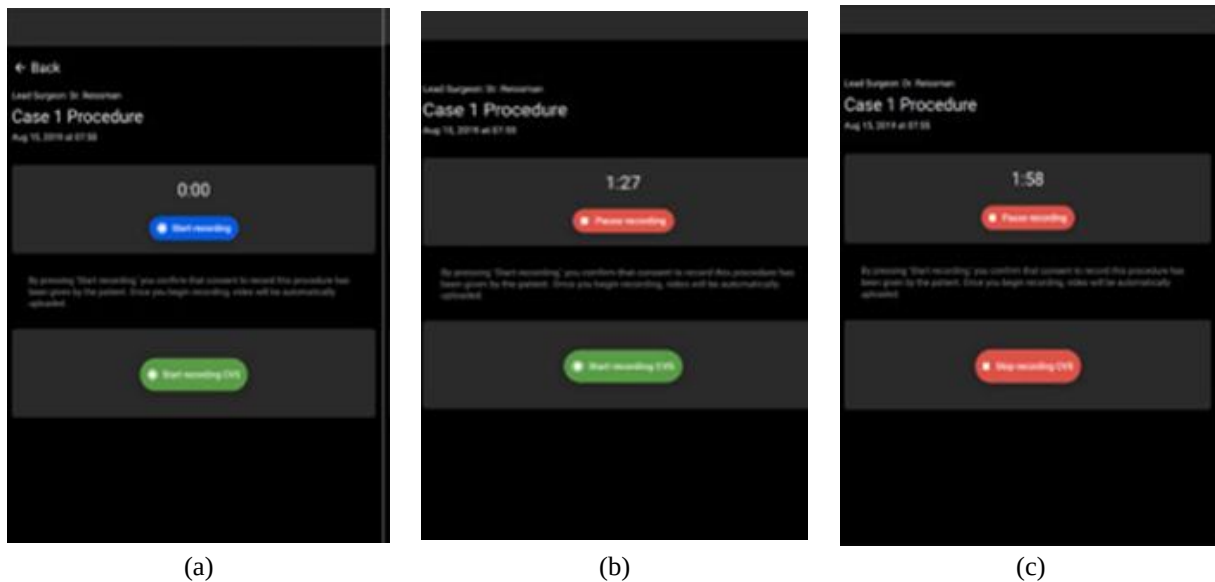


Figure 2. Operator interface displayed step by step.

Step 1: Before commencing the operation, the operator activates recording by pressing the blue “start recording”

button on the touchscreen. This initiates video capture and activates real-time processing by the AI model.

Step 2: When the surgeon confirms fulfillment of all four CVS criteria, they instruct the operator (OR nurse or assistant) to press the green “Start recording CVS” button.

Step 3: After completion of the CVS phase and upon entry of the clipping instrument into the operative field, the operator presses the “Stop recording CVS” button.

Following the outcomes of the current investigation, a subsequent phase 2 trial is planned. In this next stage, the software will provide direct visual feedback to the surgeon via a secondary monitor, clearly displaying the status of the CVS criteria.

Statistical analysis

We assessed model performance through multi-class ROC (Receiver Operating Characteristic) analysis,

calculating the AUC (Area Under the Curve) individually for each CVS criterion.

Overall CVS attainment was defined as meeting at least 3 of the 4 criteria.

We derived the 95% empirical confidence intervals (CI) for CVS achievement rates using bootstrapping across the video cases.

Results and Discussion

Rate of CVS achievement

As shown in **Table 1**, the CVS achievement rate in the general dataset (used to train and test the AI-ML model) was 18.7% [95% CI: 17.1%, 19.9%]. In contrast, during the prospective study, this rate rose substantially to 85% [95% CI: 84.4%, 85.2%].

Table 1. Rate of CVS criteria achievement (95% CI is given in brackets).

Empty cell	Training/Evaluation dataset (700 procedures)	Study group (40 procedures)
C1	34.8% [32.9%, 36.1%]	90.0% [89.6%, 90.2%]
C2	19.4% [18.1%, 20.2%]	85.0% [84.1%, 85.4%]
C3a	28.7% [27.1%, 29.5%]	87.5% [86.9%, 88.1%]
C4b	29.6% [27.9%, 31.2%]	85.0% [84.2%, 85.7%]
Overall CVS	18.7% [17.1%, 19.9%]	85.0% [84.4%, 85.2%]

Evaluation of the AI-ML Model

The effectiveness of the AI-ML model was assessed using the area under the curve (AUC) metric based on the annotated dataset. The evaluation was performed on a test set containing 140 videos, representing 20% of the complete dataset. The model achieved an overall AUC value of 0.90. When analyzed according to individual criteria, the AUC values were 0.89 for C1, 0.96 for C2, 0.88 for C3a, and 0.90 for C3b. The receiver operating characteristic (ROC) curves corresponding to each critical view of safety (CVS) criterion are presented in **Figure 3**. Additionally, the confusion matrices generated from the evaluation dataset are provided in **Table 2**.

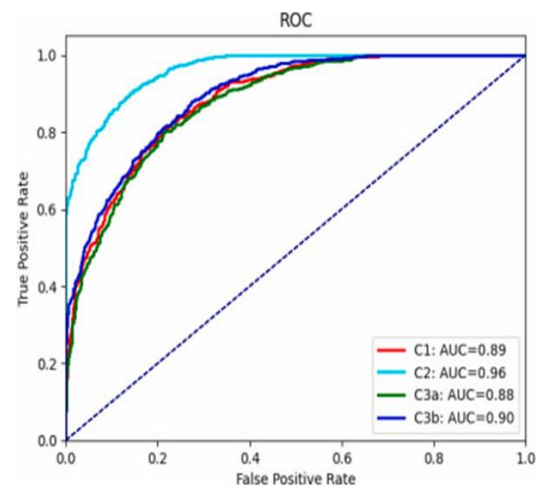


Figure 3. CVS criteria ROCs on the evaluation dataset. We evaluated the performance of our AI-ML model using ROC on our annotated dataset.

Table 2. Confusion matrices on the evaluation dataset.

C1	Actual	
Predicted	47.5 %	27.5 %
	0.7 %	49.3 %

C3b		Actual	
Predicted		32.3 %	17.7 %
		3.3 %	46.7 %
C3a		Actual	
Predicted		15.4 %	34.6 %
		0.1 %	49.9 %
C2		Actual	
Predicted		22.5 %	2.5 %
		11.6 %	38.4 %

Figure 4 presents a representative case illustrating the timing of CVS recognition by the surgeon compared to the AI-ML model. In this instance, the AI-ML model identified CVS achievement 20 seconds earlier than the surgeon. We account for the typical brief lag that occurs between the surgeon verbally announcing the milestone and actually pressing the button. Any instance where the temporal difference between the surgeon's announcement and the AI-ML model's detection remained under 30 seconds was classified as a correct decision.

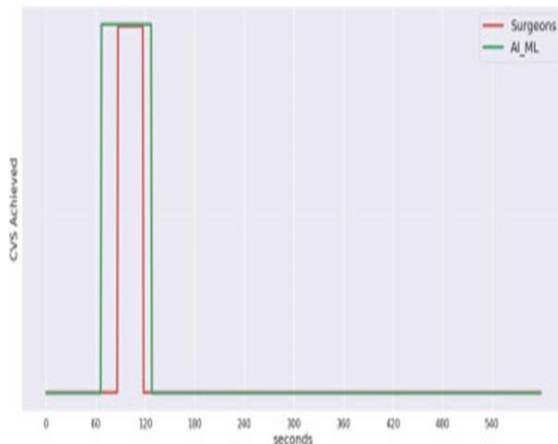


Figure 4. The timing comparison for CVS attainment between the surgeon and the AI-ML model.

The surgeon indicated that CVS was reached at 1:27 (measured from the start of the recording), whereas the AI-ML model detected it approximately 20 seconds earlier. This aligns with the procedural steps outlined in **Figure 2**.

Among the 40 laparoscopic cholecystectomy (LC) procedures, surgeons documented successful CVS in 34 cases (85%). In the remaining 6 procedures, CVS could not be achieved because of intense inflammation and significant anatomical changes. All six patients had a history of severe cholecystitis, liver abscess, or prior cholecystostomy drain placement. In every one of these

difficult cases, the AI-ML model concurred with the surgeons and correctly refrained from flagging CVS. Of the 34 procedures in which CVS was successfully achieved, the AI-ML model detected it in 33 cases, overlooking only one instance. Consequently, the AI-ML model demonstrated a specificity of 100% [95% CI: 98.1%, 100%] and a sensitivity of 97% [95% CI: 96.1%, 98.2%] across the 40 procedures evaluated in this study.

Table 3 presents the associated confusion matrix.

Table 3. Confusion matrix for the 40 procedures in the study case.

CVS	Actual	
	82.5 %	2.5 %
Predicted	0 %	15 %

Although the risk of serious bile duct injury during laparoscopic cholecystectomy (LC) is relatively low (0.3–1.5%), its occurrence has not declined over recent decades [10-15]. Given that LC ranks among the most commonly performed operations globally, including approximately one million procedures annually in the United States [14], efforts to reduce this serious complication are clearly justified. One widely accepted intraoperative approach to minimize such injuries is achieving the Critical View of Safety (CVS) [17-20]. A key investigation highlighting the importance of CVS involved video review of cases with post-LC complications and found that CVS was achieved in only 10.8% of those procedures [23]. In a separate study aimed at improving CVS rates, introducing a 5-second “time-out” to verify CVS before clipping and dividing structures raised the achievement rate from 15.9% to 44.1% [25]. Notably, none of the five patients (1.5%) who experienced bile duct injury or leak had obtained CVS.

Despite endorsement by several surgical societies—stemming from the SAGES Safe Cholecystectomy Task Force and related consensus recommendations advocating CVS as the standard method for confirming anatomy before irreversible steps [20]—reported CVS achievement rates in LC remain low, ranging from 7.2% to 51.8% across numerous publications [4, 17, 23-25]. The present study assessed an AI system previously trained to accurately recognize LC phases, with specific emphasis on CVS, across videos of varying difficulty [9]. The system was connected to the laparoscopic video feed in the operating room and operated silently in the

background while the surgeon, unaware of its output, performed the procedure. Results showed that the AI model aligned with the surgeon's assessment of CVS achievement in 97% of cases where CVS was attainable. In procedures where CVS could not be obtained due to severe local inflammation and anatomical distortion, the AI model demonstrated 100% accuracy in correctly predicting its absence.

Additionally, we observed that surgeons' awareness of an AI system monitoring for CVS in real time was associated with a noticeably higher rate of CVS achievement compared to rates documented in prior literature [4, 17, 23-25]. In this study, CVS was attained in 85% of cases—substantially above most published figures and markedly higher than the 18.7% rate observed in our own training dataset of 700 multi-source annotated LC videos.

Although the study design did not include a direct comparison of individual surgeons' CVS rates with and without the AI system (as this was not its objective), the 85% achievement rate remains noteworthy.

It appears that simply linking the AI-ML system to the operating room video feed for subsequent analysis and comparison with the surgeon's documentation was sufficient to drive a meaningful increase in CVS attainment. This unanticipated yet understandable outcome emerged in the initial phase of a broader investigation, in which the AI will eventually deliver real-time visual guidance to the surgeon during LC.

Beyond the AI system's high accuracy in matching surgeons' CVS evaluations, its presence during live surgery—combined with surgeons' awareness of its monitoring role—triggered the well-documented Hawthorne effect [26]. Surgeons appeared to improve their performance when they knew the AI was observing their actions.

We anticipate that, following further large-scale validation, such AI systems could become standard components of operating room video platforms for LC and, eventually, more complex operations. Additional promising applications arising from this work include objective skill assessment and training progression tracking for surgical residents, using parameters such as time to CVS, CVS success rate, instrument exchanges, and other metrics. We also intend to extend this approach to more advanced laparoscopic procedures, such as adrenalectomy and pancreatectomy, by training the AI-ML model to recognize their critical phases in a similar stepwise manner.

Conclusion

In conclusion, a well-trained AI-ML model can reliably evaluate CVS during LC in a real-time operating room environment. Moreover, employing such a system intraoperatively may increase CVS attainment rates, thereby enhancing procedural safety and patient outcomes. These results support the need for additional research on integrating AI tools with laparoscopic operating room systems.

Limitations

This feasibility study involved a modest number of patients and examined the performance of an AI system running in parallel with the operating surgeon during LC. It was not designed as a randomized controlled trial. A larger prospective randomized study is planned to assess the clinical impact of AI assistance during LC more thoroughly.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Maier-Hein L, Vedula SS, Speidel S, Navab N, Kikinis R, Park A, et al. Surgical data science for next-generation interventions. *Nat Biomed Eng.* 2017;1(9):691-6.
2. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature.* 2020;577(7788):89-94.
3. Berzin TM, Topol EJ. Adding artificial intelligence to gastrointestinal endoscopy. *Lancet.* 2020;395(10223):485.
4. Madani A, Namazi B, Altieri MS, Hashimoto DA, Rivera AM, Pucher PH, et al. Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. *Ann Surg.* 2020; [Online ahead of print].
5. Mascagni P, Vardazaryan A, Alapatt D, Urade T, Felli E, Pessaux P, et al. Artificial intelligence for

- surgical safety: automatic assessment of the critical view of safety in laparoscopic cholecystectomy using deep learning. *Ann Surg.* 2022;275(5):955-61.
6. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg.* 2018;268(1):70-6.
 7. Hung AJ, Chen J, Gill IS. Automated performance metrics and machine learning algorithms to measure surgeon performance and anticipate clinical outcomes in robotic surgery. *JAMA Surg.* 2018;153(8):770-1.
 8. Twinanda A, Shehata S, Mutter D, Marescaux J, de Mathelin M, Padoy N. EndoNet: a deep architecture for recognition tasks on laparoscopic videos. *IEEE Trans Med Imag.* 2017;36(1):86-97.
 9. Golany T, Aides A, Freedman D, Rabin N, Sela Y, Zuroff A, et al. Artificial intelligence for phase recognition in complex laparoscopic cholecystectomy. *Surg Endosc.* 2022;36(12):9215-23.
 10. Schwaitzberg SD, Scott DJ, Jones DB, McKinley SK, Sillin L, Brunt LM. Threefold increased bile duct injury rate is associated with less surgeon experience in an insurance claims database. *Surg Endosc.* 2014;28:3068-73.
 11. Törnqvist B, Strömberg C, Akre O, Enochsson L, Nilsson M. Selective intraoperative cholangiography and risk of bile duct injury during cholecystectomy. *Br J Surg.* 2015;102:952-8.
 12. Lilley EJ, Scott JW, Jiang W, Thomas J, Cooper Z, Barash D, et al. Intraoperative cholangiography during cholecystectomy among hospitalized medicare beneficiaries with non-neoplastic biliary disease. *Am J Surg.* 2017;214:682-6.
 13. Barrett M, Asbun HJ, Chien HL, Brunt LM, Telem DA. Bile duct injury and morbidity following cholecystectomy: a need for improvement. *Surg Endosc.* 2018;32:1683-8.
 14. Pucher PH, Brunt LM, Fanelli RD, Asbun HJ, Aggarwal R. Outcome trends and safety measures after 30 years of laparoscopic cholecystectomy: a systematic review and pooled data analysis. *Surg Endosc.* 2018;32:2175-83.
 15. Fong ZV, Pitt HA, Strasberg SM, Loehrer AP, Chang DC, Win C, et al. Diminished survival in patients with bile leak and ductal injury: management strategy and outcomes. *J Am Coll Surg.* 2018;226:568-76.
 16. Way LW, Stewart L, Gantert W, Liu K, Lee CM, Whang K, et al. Causes and prevention of laparoscopic bile duct injuries: analysis of 252 cases from a human factors and cognitive psychology perspective. *Ann Surg.* 2003;237(4):460-9.
 17. Strasberg SM, Brunt LM. Rationale and use of the critical view of safety in laparoscopic cholecystectomy. *J Am Coll Surg.* 2010;211:132-8.
 18. Avgerinos C, Kelgiorgi D, Touloumis Z, Baltatzis M, Derveniz C. One thousand laparoscopic cholecystectomies in a single surgical unit using the "critical view of safety" technique. *J Gastrointest Surg.* 2009;13(3):498-503.
 19. Manatakis DK, Papageorgiou D, Antonopoulou MI, Kyriazanos ID, Derveniz C, Koutelidakis I, et al. Ten-year audit of safe bail-out alternatives to the critical view of safety in laparoscopic cholecystectomy. *World J Surg.* 2019;43(11):2728-33.
 20. Brunt LM, Deziel DJ, Telem DA, Strasberg SM, Aggarwal R, Asbun HJ, et al. Safe cholecystectomy multi-society practice guideline and state-of-the-art consensus conference on prevention of bile duct injury during cholecystectomy. *Surg Endosc.* 2020;34:2827-55. (co-publication: *Ann Surg.* 2020;272:3-23)
 21. Du X, Lin T, Jin P, Chen G, Liu X, Tan M, et al. SpineNet: learning scale-permuted backbone for recognition and localization. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2020.
 22. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016.
 23. Nijssen MA, Shreimakers JM, Meyer Z, van der Voort M, Bonjer HJ, Eijssbouts QAJ. Complications after laparoscopic cholecystectomy: a video evaluation study of whether the critical view of safety was reached. *World J Surg.* 2015;39:1798-803.
 24. Mascagni P, Rodríguez-Luna MR, Urade T, Felli E, Pessaux P, Mutter D, 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.* 2021;233(4):497-505.
 25. Stefanidis D, Chintalapudi N, Anderson-Montoya B, Oza P, Sarker S. How often do surgeons obtain the

- critical view of safety during laparoscopic cholecystectomy? *Surg Endosc.* 2017;31:142-6.
26. Liu R, Liu Z, Wang X. Intraoperative time-out to promote the implementation of the critical view of safety in laparoscopic cholecystectomy: on the comprehension of criteria, disease severity, and Hawthorne effects. *J Am Coll Surg.* 2022;34(6):1261-2.
 27. Owen D, Grammatikopoulou M, Luengo I, Stoyanov D. Detection of critical structures in laparoscopic cholecystectomy using label relaxation and self-supervision. In: *MICCAI Proceedings.* 2021;24(IV):321-30.
 28. Namazi B, Iyengar N, Hasan S, Balachandra S, Madani A, Hashimoto D, et al. AI for automated detection of the establishment of Critical View of Safety in laparoscopic cholecystectomy videos. *J Am Coll Surg.* 2020;231(4):e48.
 29. Colbeci Y, Zohar M, Neimark D, Asselmann D, Bar O. A multi instance learning approach for critical view of safety detection in laparoscopic cholecystectomy. In: *Proceedings of the 7th Machine Learning for Healthcare Conference.* PMLR; 2022;182:409-24.
 30. Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, et al. Microsoft COCO: common objects in context. In: *13th European Conference on Computer Vision (ECCV) Proceedings.* 2014;13:740-55.