



Development of a deep learning model for safe direct optical trocar insertion in minimally invasive surgery: an innovative method to prevent trocar injuries

Supakool Jearanai¹ · Piyanun Wangkulangkul¹ · Wannipa Sae-Lim² · Siripong Cheewatanakornkul¹

Received: 31 May 2023 / Accepted: 12 July 2023 / Published online: 9 August 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Background Direct optical trocar insertion is a common procedure in laparoscopic minimally invasive surgery. However, misinterpretations of the abdominal wall anatomy can lead to severe complications. Artificial intelligence has shown promise in surgical endoscopy, particularly in the employment of deep learning models for anatomical landmark identification. This study aimed to integrate a deep learning model with an alarm system algorithm for the precise detection of abdominal wall layers during trocar placement.

Method Annotated bounding boxes and assigned classes were based on the six layers of the abdominal wall: subcutaneous, anterior rectus sheath, rectus muscle, posterior rectus sheath, peritoneum, and abdominal cavity. The cutting-edge YOLOv8 model was combined with a deep learning detector to train the dataset. The model was trained on still images and inferenced on laparoscopic videos to ensure real-time detection in the operating room. The alarm system was activated upon recognizing the peritoneum and abdominal cavity layers. We assessed the model's performance using mean average precision (mAP), precision, and recall metrics.

Results A total of 3600 images were captured from 89 laparoscopic video cases. The proposed model was trained on 3000 images, validated with a set of 200 images, and tested on a separate set of 400 images. The results from the test set were 95.8% mAP, 89.8% precision, and 91.7% recall. The alarm system was validated and accepted by experienced surgeons at our institute.

Conclusion We demonstrated that deep learning has the potential to assist surgeons during direct optical trocar insertion. During trocar insertion, the proposed model promptly detects precise landmark references in real-time. The integration of this model with the alarm system enables timely reminders for surgeons to tilt the scope accordingly. Consequently, the implementation of the framework provides the potential to mitigate complications associated with direct optical trocar placement, thereby enhancing surgical safety and outcomes.

Keywords Abdominal wall detection · Deep learning · Direct optical trocar · Minimally invasive surgery · Laparoscopy

Trocar insertion is a critical step in laparoscopic surgery with various methods available for introducing a trocar into the abdominal cavity. Direct optical trocars are widely used to establish pneumoperitoneum in most patients who

undergo bariatric surgery. Therefore, direct optical trocars are particularly suitable for patients with a high body mass index (BMI) who require minimally invasive procedures. Compared to the open (Hasson) technique, the direct optical trocar method reduces the time needed, blood loss, and incidence of abdominal organ injury in obese patients [1, 2]. However, this procedure carries the risk of severe complications such as hollow viscus organ rupture or catastrophic aortic rupture [3–5].

In recent years, deep learning artificial intelligence (AI) has demonstrated significant potential for various medical applications including the enhancement of surgical procedures [6, 7]. Several studies have explored the use of deep learning

✉ Siripong Cheewatanakornkul
siripong.c@psu.ac.th

¹ Minimally Invasive Surgery Unit, Department of Surgery, Faculty of Medicine, Prince of Songkla University, Songkhla, Thailand

² Department of Computer Science, Faculty of Science, Prince of Songkla University, Songkhla, Thailand

techniques, which is a subset of AI, for different aspects of laparoscopic surgery. Surgical phase recognition is one of the essential aspects of laparoscopic surgery because it helps surgeons monitor the steps of surgery and enables efficient surgical workflow management [8]. Deep learning models have shown success in automatic surgical phase recognition in laparoscopic procedures. In 2019, Padoy et al. [9] used a convolutional neural network (CNN)-based approach to recognize surgical phases from video data that demonstrated the ability of the model to learn complex spatiotemporal patterns. More recently, Yu et al. proposed a deep learning framework for surgical phase recognition using a combination of CNNs and recurrent neural networks (RNNs) [10] and achieved high accuracy in predicting different surgical phases. Furthermore, surgical instrument recognition is another important task for laparoscopic surgery as it can track the surgical tools in real-time. Various deep learning techniques have been employed to detect and classify surgical instruments in laparoscopic images. For instance, García-Peraza-Herrera et al. [11] used a CNN-based approach for real-time surgical instrument detection while Jha et al. [12] proposed a deep learning model for instrument segmentation in laparoscopic videos to achieve high accuracy and real-time performance. In the field of anatomical landmark detection, deep learning-based approaches have been employed for safe and accurate surgical navigation. Many studies proposed real-time identification of important anatomical landmarks to guide surgeons during laparoscopy. Twinanda et al. [13] employed a CNN to detect and segment anatomical landmarks in laparoscopic images. They demonstrated improved performance compared to traditional image processing techniques. Similarly, Wang et al. [14] proposed a deep learning model for the automatic detection of anatomical landmarks in endoscopic images to achieve high accuracy and precision.

In this study, we aimed to integrate a deep learning approach with a laparoscopic suite equipped with an alarm system that would enable surgeons to identify safety landmarks in real-time that would potentially increase operation safety. Furthermore, the details of the development and evaluation of the deep learning framework are described for automatic abdominal wall detection during trocar insertion. We outline the methodology employed to train the YOLOv8 model, demonstrate its performance in detecting the six layers of the abdominal wall, compare it with other YOLO models, and discuss evaluation of the alarm system.

Materials and methods

Surgical technique

The traditional direct optical trocar insertion technique allows surgeons to clearly visualize the abdominal wall

anatomy as outlined in the following steps. The patient is in the supine position.

Step 1: Incisions are made above the rectus muscle area.

Step 2: A 12 mm 0° laparoscopic camera is attached to the trocar.

Step 3: The surgeon performs the procedure by applying perpendicular pressure while inserting the instruments through the layers of the abdominal wall that include the subcutaneous, anterior rectus sheath, rectus muscle, posterior rectus sheath, and peritoneum.

Step 4: Upon observing peritoneal tear patterns, the surgeon tilts the camera to position it parallel to the patient's abdominal wall before continuing with insertion of the remaining portion of the scope into the abdominal cavity.

Step 5: Pneumoperitoneum is established through CO₂ insufflation.

Data collection and annotation

We retrospectively collected 89 laparoscopic trocar placement videos at Songklanagarind Hospital, which is a teaching hospital on the campus of Prince of Songkla University. These videos were collected from different patient demographics, various camera models, and camera angles. To ensure ethical research conduct, all identifiable patient information was deidentified and anonymized. A team of three surgeons extracted 3600 still images of the trocar insertion procedure in the videos. Annotations were then made to establish landmark bounding boxes and classify object classes according to the six layers of the abdominal wall (Fig. 1). These layers included the subcutaneous, anterior rectus sheath, rectus muscle, posterior rectus sheath, peritoneum, and abdominal cavity. We utilized the Roboflow [15] web application, which is an open-source tool, for the annotation process. The collection of retrospective medical data and the subsequent analysis were conducted in accordance with the ethical standards set forth by the institutional and national research committee as well as the 2013 version of the Declaration of Helsinki (REC Approval Number 65-161-10-3).

Data splitting

The complete dataset consisted of 3600 images separated into three subsets: a training set of 3000 images, a validation set of 200 images, and a test set of 400 images.

Data pre-processing

To train the detection model, we resized all images in the dataset from 3840 × 2160 pixels to 800 × 800 pixels using the resize function of the YOLOv8 model. Subsequently,

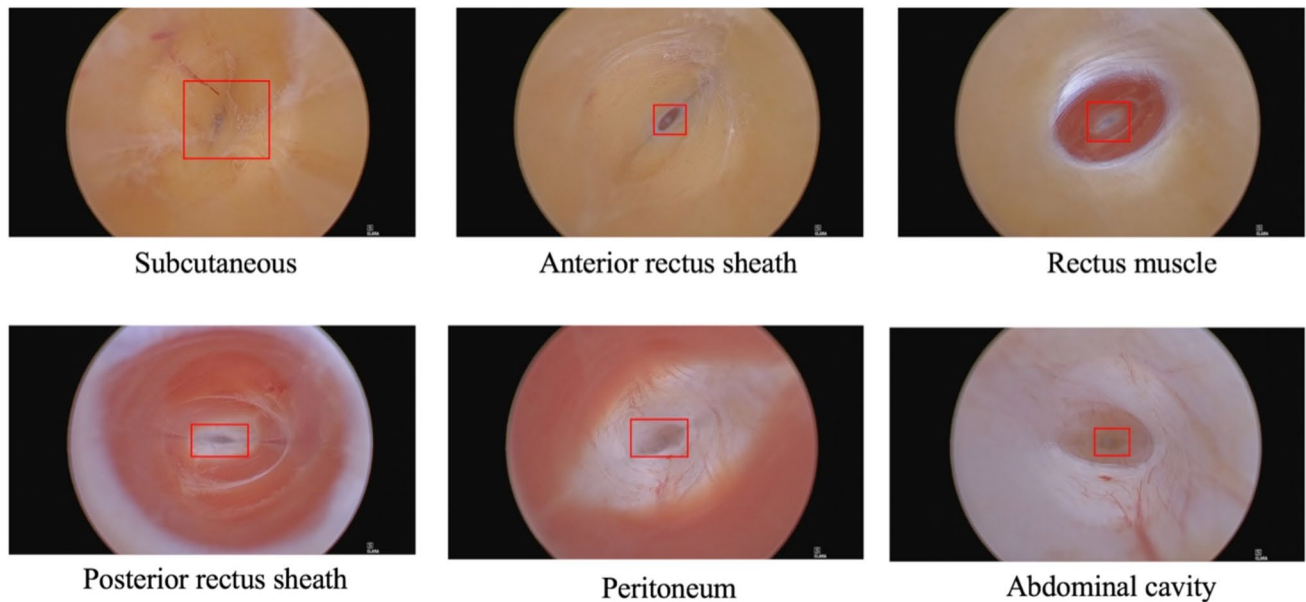


Fig. 1 Example annotations of the abdominal wall layer

the training set was augmented by the Mosaic augmentation approach [7], which provided additional images at various scales, orientations, contrasts, and brightness levels. Mosaic augmentation worked by randomly selecting four images and combined them into a single image. Given the chosen batch size of 16, the resultant four images were resized and arranged in a 4×4 grid to create a mosaic pattern. This configuration allowed the model to simultaneously train on 64 images during each iteration. Mosaic augmentation enabled the model to learn each layer of the abdominal wall in various appearances and contexts, which therefore enhanced the prediction accuracy in real-world laparoscopic scenarios and prevented overfitting during the training process [16]. Figure 2 presents the 4×4 grid of 16 batch size that resulted from Mosaic augmentation.

YOLOv8 model

YOLOv8 (You Only Look Once version 8) is a state-of-the-art, real-time object detection model developed by Ultralytics [17]. YOLOv8 was constructed as a single-stage detector using a deep CNN to simultaneously predict bounding boxes and class probabilities for multiple objects. The architecture of YOLOv8 consists of a backbone network designed to extract features from input images and a head network that predicts bounding boxes and the object classes (Fig. 3). Within the backbone and head network, YOLOv8 incorporates multiple layers of convolution, pooling, and combining along with batch normalization and activation functions. The overall architecture of YOLOv8 was developed based on previous versions

of the YOLO family that resulted in several improvements. The performance of YOLOv8 was enhanced by incorporating various techniques, such as Mosaic augmentation, that generate additional training images by randomly combining four images into one. Additionally, YOLOv8 employs focal loss to address class imbalance to optimize the detection of difficult objects.

Model training and fine-tuning

We employed the YOLOv8 model for dataset training. The training model consisted of 225 layers with 11,137,922 parameters. The training process was carried out in two stages. In the first stage, the model was trained for 200 epochs using the pre-trained weights provided by the default package of the model. This initial stage aimed to adapt the model to the specific dataset and achieve a reasonable level of performance. Upon evaluating the prediction performance after the initial stage, we proceeded to refine the hyperparameters, re-annotate the dataset, and address class imbalance. The second stage of fine-tuning was performed for an additional 100 epochs using the weights obtained from the first stage. This step aimed to refine the model further, which allowed it to achieve higher accuracy and better generalization for the task of abdominal wall detection. All training processes were conducted on the Google Colab platform utilizing the NVIDIA A100 graphics processing unit. Various hyperparameters were optimized to achieve the optimum performance of the model (Table 1).

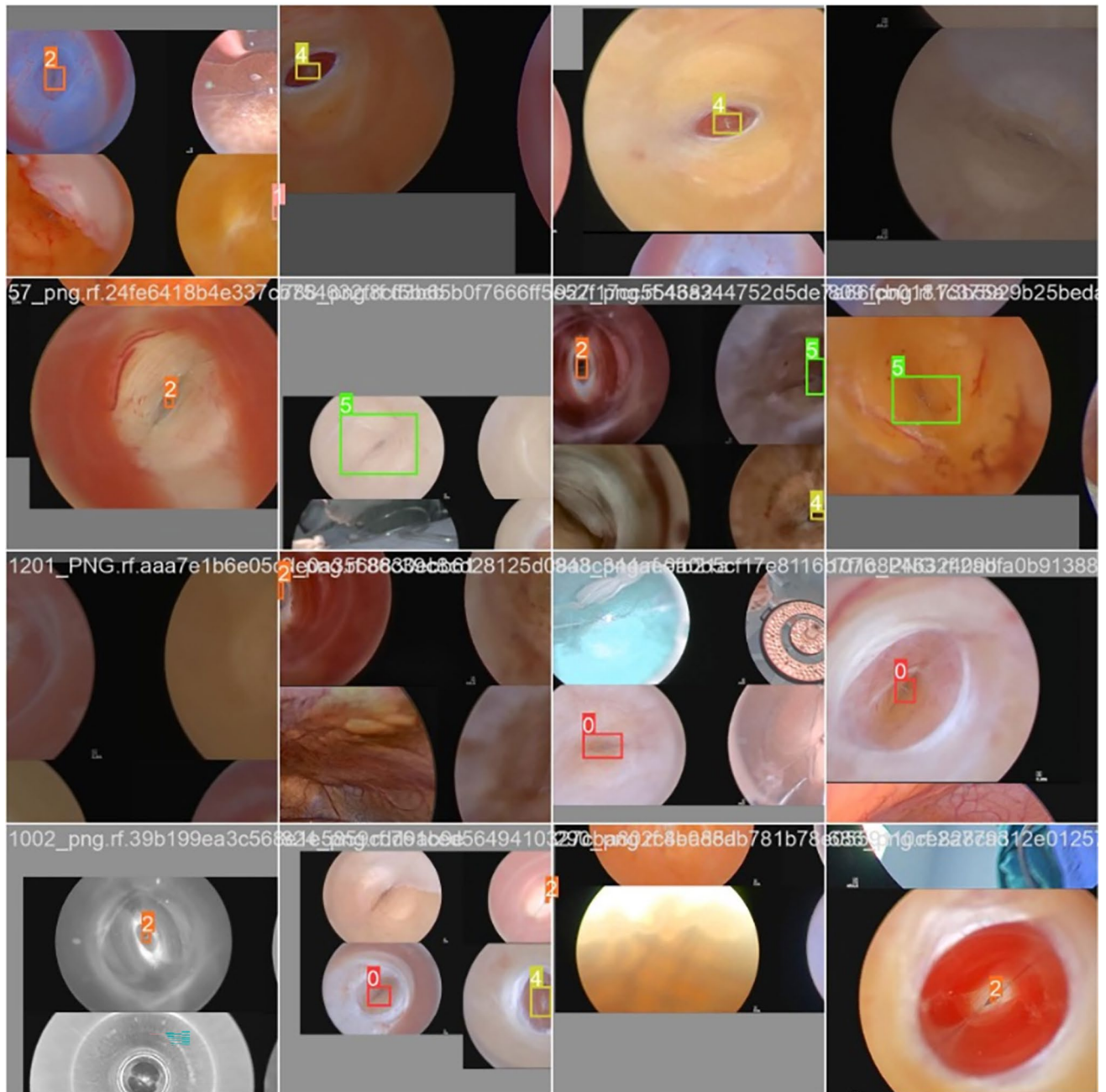


Fig. 2 Mosaic augmentation results of 16 batch size

Alarm sound algorithm

An alarm system was developed to provide notifications during trocar placement through the peritoneum into the abdominal cavity. We integrated this algorithm with the YOLOv8 detection process to sound the first alarm when the predicted bounding box for the peritoneum appeared. The second alarm was activated when the predicted bounding box for the abdominal cavity became visible.

Model testing and evaluation

The performance of the detection model was assessed using unseen images in the test set. Precision, recall, and average precision (AP) were utilized as individual class evaluations. Furthermore, the overall performance of the model was quantified using mean average precision (mAP), which denoted the mean value of all APs across classes. To achieve the mAP value, the first step was to calculate the intersection

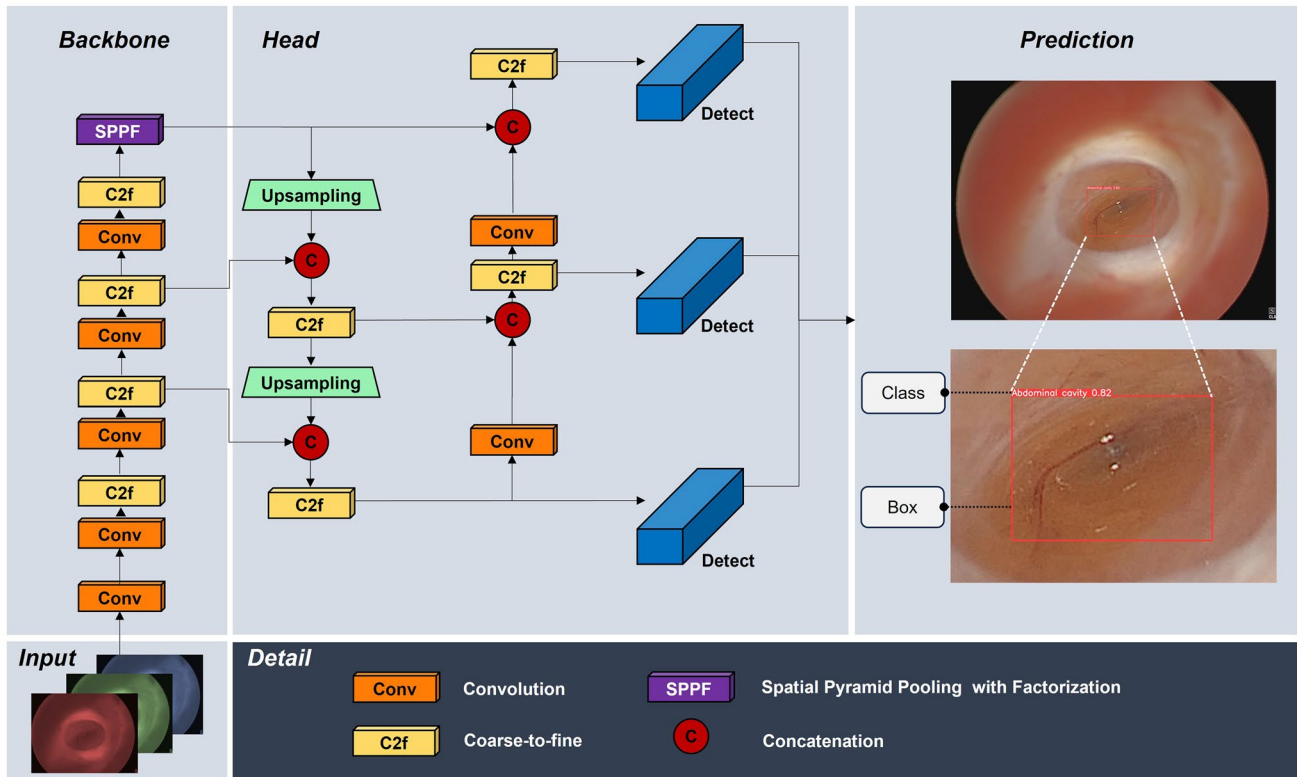


Fig. 3 Architecture of proposed framework based on the YOLOv8 model

Table 1 Hyperparameter details for the training process

Hyperparameter	First stage	Second stage (fine-tuning)
Epoch	200	100
Batch size	16	16
Image size (pixels)	800 × 800	640 × 640
Pre-trained weight	Yes	Transferred from first stage
Optimizer	Stochastic gradient descent	Stochastic gradient descent
Initial learning rate	0.001	0.01
Final learning rate	0.01	0.01
Momentum	0.937	0.937
Albumentations	Blur, MedianBlur, ToGray, CLAHE	HSV, Flip, Scale, Blur, MedianBlur, ToGray, CLAHE

HSV hue saturation value

over union (IoU) as depicted in Fig. 4 (Eq. 1), where BB_{pred} represents the predicted bounding box and BB_{gt} represents the ground truth bounding box. The IoU metric denotes the proportion of the intersecting region to the combined area of the predicted bounding box and the ground truth bounding box (Fig. 4). Subsequently true positive, false positive, and false negative were calculated by comparing the

IoU values with the predefined thresholds. For each abdominal wall layer, precision and recall values as well as their associated confidence levels were determined using Fig. 4 (Eqs. 2 and 3). Next, the AP for a specific class was calculated using Fig. 4 (Eq. 4) where n represents the number of threshold points, i represents the index of each threshold, P_i refers to the precision values, and R_i corresponds to the recall values. Lastly, the mAP was calculated by taking the mean of the AP values for each class as presented in Fig. 4 (Eq. 5) where AP_k refers to the AP of class k , and n refers to the number of classes.

Results

Detection performance of the six layers of the abdominal wall

We trained the YOLOv8 model for 200 epochs on the dataset and subsequently fine-tuned it for an additional 100 epochs using the re-annotated dataset. The performance of the model for each abdominal wall layer was assessed using precision, recall, and AP while the mAP measured the overall model performance. Figure 4 Eqs. 1 through 5 were employed for the evaluation calculations. The fine-tuned

Fig. 4 Intersection over union metric

$$IoU = \frac{\text{area of intersection}}{\text{area of union}} = \frac{BBox_{pred} \cap BBox_{gt}}{BBox_{pred} \cup BBox_{gt}} \quad (1)$$

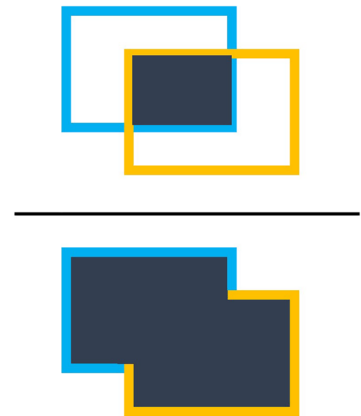
$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$AP = \sum_{i=0}^{n-1} P_i(R_i - R_{i+1}) \quad (4)$$

$$mAP = \frac{1}{n} \sum_{k=1}^n AP_k \quad (5)$$

Intersection over Union (IoU) =



model achieved an overall detection performance of 89.8% precision, 91.7% recall, and 95.8% mAP on the testing set. Table 2 presents the precision, recall, and AP results for each abdominal wall layer of the testing sets. Figure 5 illustrates the abdominal wall detection results on the testing set.

Alarm system performance

We inferenced the well-trained detection model during the detection process to 15 laparoscope videos to evaluate the performance of the alarm sound algorithm. Nine minimally invasive surgeons verified the accuracy of the alert sound timing and duration. Table 3 lists the number of acceptances of the alarm sound algorithm. The results in Table 3 indicate that the alert sound was activated correctly during the trocar

insertion process. The first sound was produced properly when the result of the predicted bounding box was the peritoneum layer and the second sound was subsequently heard when the abdominal cavity was detected.

Table 2 Evaluation of the abdominal wall detection model for all classes of the testing set with an intersection over union threshold of 0.5

Class	Precision (%)	Recall (%)	AP (%)
Subcutaneous	83.9	94.9	96.8
Anterior rectus sheath	82.8	92.7	97.4
Rectus muscle	90.9	95.5	98.0
Posterior rectus sheath	93.8	87.8	94.5
Peritoneum	96.3	95.4	96.9
Abdominal cavity	91.1	84.1	91.0
Average	89.8	91.7	95.8 (mAP)

AP average precision, mAP mean average precision

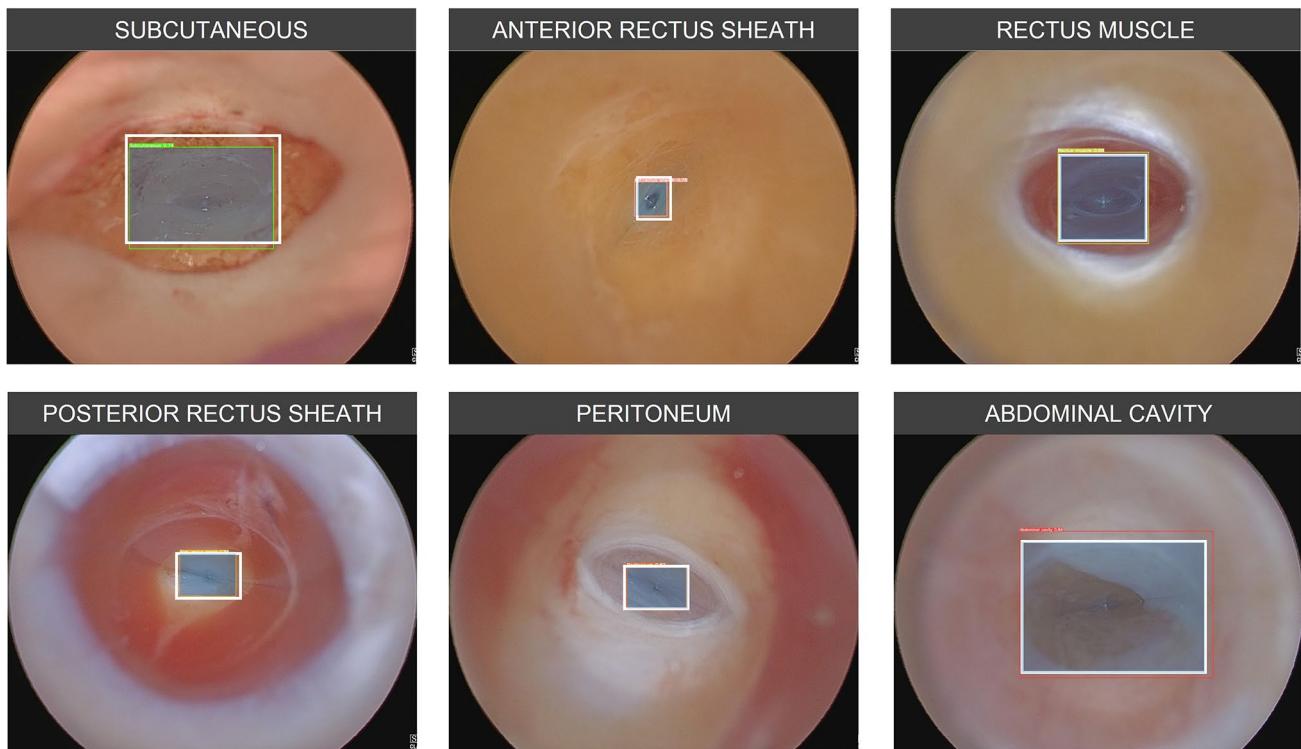


Fig. 5 Proposed model prediction compared to ground truth annotation

Table 3 Surgeons' acceptance of the alarm sound algorithm on the testing set

Surgeon No	Number of acceptances and rejections	
	Accepted	Rejected
1	15	0
2	15	0
3	15	0
4	14	1
5	14	1
6	15	0
7	15	0
8	15	0
9	14	1
Average (%)	97%	3%

Comparisons with scaled-YOLOv4, YOLOv5, and YOLOv7

While scaled-YOLOv4, YOLOv5, and YOLOv7 have been employed as robust object detection models, the redesigned architecture of YOLOv8 has proven to be more effective in terms of both speed and accuracy. The detection

Table 4 Comparison of performance on the testing set

Model	Precision (%)	Recall (%)	mAP (IoU=0.5) (%)
Scaled-YOLOv4	62.5	74.7	70.7
YOLOv5	64.3	66.7	62.3
YOLOv7	70.2	79.1	76.0
YOLOv8	89.8	91.7	95.8

mAP mean average precision, *IoU* intersection over union

performance of the YOLOv8 model was compared against the scaled-YOLOv4, YOLOv5, and YOLOv7 models on the testing set. The mAP of each model was calculated with an IoU threshold of 0.5. Table 4 presents the comparison results for precision, recall, and mAP. Based on all values, the findings given in Table 4 demonstrated that the YOLOv8 model outperformed the scaled-YOLOv4, YOLOv5, and YOLOv7 models.

Real-time inference in the operating room

The following steps were undertaken in the operating room to evaluate the framework for real-time abdominal wall detection during trocar insertion.

Step 1: The necessary hardware and software were made available: (1) a compatible endoscopic camera system with 4 K resolution or the highest available resolution from the camera control unit, (2) a computer with Intel Core i7 processor, 16 GB of RAM, and NVIDIA GeForce RTX 3080 graphics card for real-time inference, and (3) the well-trained YOLOv8 model including the configuration file, weight file, and object class file.

Step 2: The operating room set-up was performed by: (1) connecting the endoscopic camera system to the camera control unit (CCU) and (2) transmitting the live video from the CCU to the computer for real-time processing where the HDMI capture card was used to convert the endoscopic video signal to digital data that was fed into the framework.

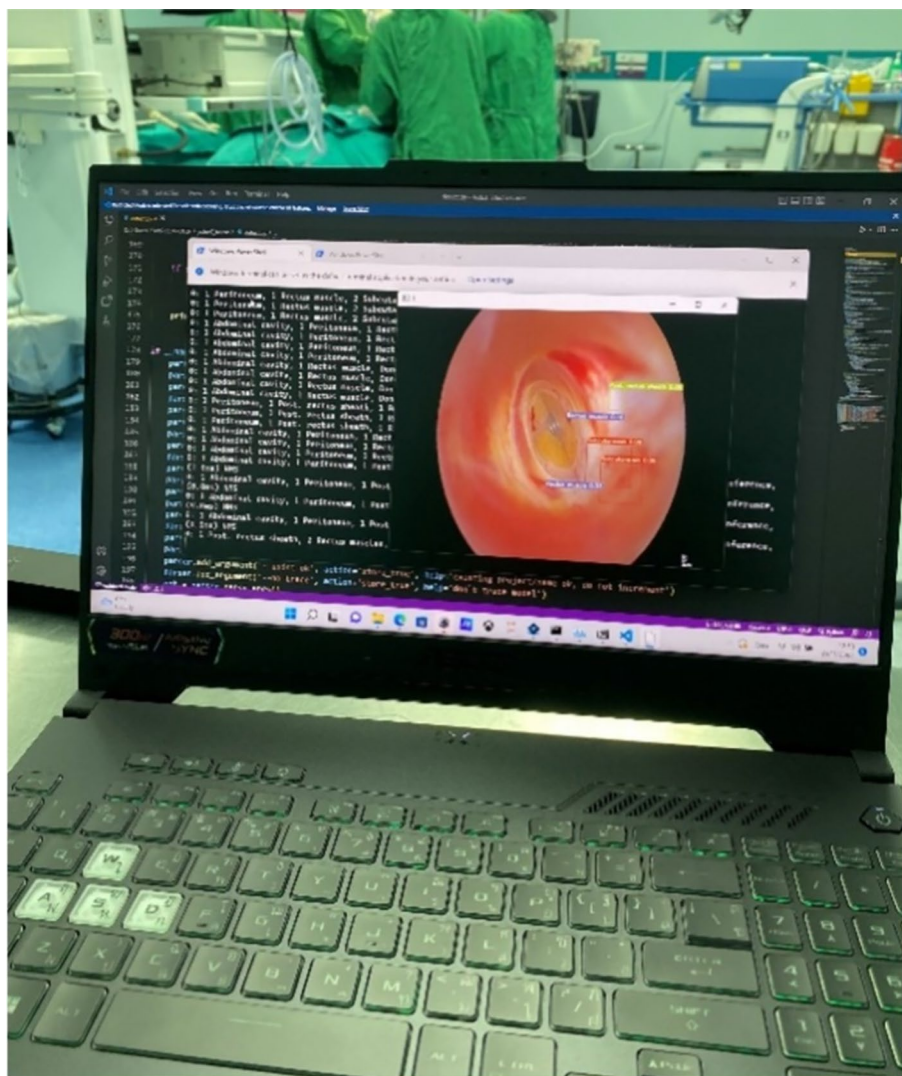
Step 3: The framework was loaded for real-time inference by: (1) configuring the framework to process the live

video from the endoscopic camera system, (2) initiating real-time inference using the well-trained YOLOv8 model that enabled it to detect the abdominal wall layers as the trocar was inserted, and (3) calling the alarm function to notify the surgeon when the trocar was inserted into the peritoneum layer and abdominal cavity layer.

Step 4: The results were monitored. As the surgeon performed the trocar insertion, the real-time output was monitored, the prediction and activation of the alarm function were recorded, and precision was evaluated by the surgeon team.

By following these steps, the well-trained model detected each layer of the abdominal wall and accurately performed the alarm system. Figure 6 presents the workstation in the operating room.

Fig. 6 Workstation in the operating room



Discussion

Several methods are available for entering the abdominal cavity in minimally invasive surgery. Direct optical trocar insertion is the prevalent surgical technique, particularly for obese or high BMI patients. However, this technique can raise concerns for surgeons, as improper use may lead to severe complications. This study employed AI to identify abdominal wall entry patterns and develop an alarm system to help surgeons prevent intra-abdominal organ damage. This system is also suitable for training inexperienced surgeons in using this method. The test results demonstrated that the AI system developed at our institution can accurately recognize abdominal wall anatomy and provide alerts when optimal peritoneal tearing is achieved.

AI is progressively being employed across various facets of surgery to enhance patient outcomes, optimize workflows, and improve surgical training [18]. In the domain of intraoperative guidance and navigation, AI can be integrated with surgical robots or computer-assisted systems to provide real-time guidance and feedback during surgery. This integration was shown to bolster accuracy and efficiency in surgical procedures such as laparoscopic cholecystectomy where it assists in identifying critical safety views and in colorectal surgery [8] where it aids in pinpointing vital structures.

Trocar insertion during the laparoscopic procedure is recognized at our institution as a crucial factor in patient safety. In this study, we successfully implemented a YOLOv8-based deep learning model for real-time abdominal wall detection during direct optical trocar insertion. The findings demonstrated the potential of this approach to substantially improve the safety and accuracy of trocar placement.

The deep learning framework allowed for precise identification and localization of the six layers of the abdominal wall. This real-time guidance enables surgeons to be constantly aware of the anatomical structures encountered during the procedure, which subsequently reduces the risk of trocar injuries and associated complications. Furthermore, the integration of an alarm system within the model ensures timely alerts for the surgeon to enhance surgical decision-making and facilitate adjustments as needed.

This study had some limitations. It was a single-center study. The surgical videos used for training were sourced exclusively from our hospital's laparoscopic equipment. Some equipment brands were not included in the development of this model, which may require additional sampling and training to enhance its adaptability. Furthermore, future randomized controlled trials should be conducted to investigate the benefits of implementing AI in surgical procedures.

In conclusion, deep learning technology for real-time abdominal wall detection during trocar insertion in laparoscopic surgery offers a promising method to improve patient

safety and preventing trocar-related injuries. The YOLOv8-based model presented in this study represents a significant advancement in the field of AI in surgical endoscopy with the potential to transform clinical practice and ultimately enhance surgical outcomes.

Acknowledgements We would like to thank the surgical team at Songklanagarind Hospital for their invaluable assistance in establishing the project set-up within the operating room and for their insightful feedback during laparoscopic surgery.

Author contributions All authors contributed to the drafting of the manuscript.

Funding No funding was provided for this study.

Declarations

Disclosures Supakool Jearanai, Siripong Cheewatanakornkul, Piyanun Wangkulangkul, and Wannipa Sae-Lim affirm that they have no conflicts of interest to disclose.

References

1. Tinelli A, Malvasi A, Mynbaev OA, Tsin DA, Davila F, Dominguez G et al (2013) Bladeless direct optical trocar insertion in laparoscopic procedures on the obese patient. *J Soc Laparoendosc Surg* 17(4):521–528
2. Ambardar S, Cabot J, Cekic V, Baxter K, Arnell TD, Forde KA et al (2009) Abdominal wall dimensions and umbilical position vary widely with BMI and should be taken into account when choosing port locations. *Surg Endosc* 23(9):1995–2000
3. Usman R, Ahmed H, Ahmed Z, Ali M (2020) Optical trocar causing aortic injury: a potentially fatal complication of minimal access surgery. *J Coll Physicians Surg* 30(1):85–87
4. Sharp HT, Dodson MK, Draper ML, Watts DA, Doucette RC, Hurd WW (2002) Complications associated with optical-access laparoscopic trocars. *Obstet Gynecol Surv* 57(8):502–503
5. Sundbom M, Hedberg J, Wanhainen A, Ottosson J (2014) Aortic injuries during laparoscopic gastric bypass for morbid obesity in Sweden 2009–2010: a nationwide survey. *Surg Obes Relat Dis* 10(2):203–207
6. Rajkomar A, Dean J, Kohane I (2019) Machine learning in medicine. *N Engl J Med* 380(14):1347–1358
7. Kitaguchi D, Takeshita N, Hasegawa H, Ito M (2021) Artificial intelligence-based computer vision in surgery: recent advances and future perspectives. *Ann Gastroenterol Surg* 6(1):29–36
8. Kitaguchi D, Takeshita N, Matsuzaki H, Hasegawa H, Igaki T, Oda T et al (2022) Deep learning-based automatic surgical step recognition in intraoperative videos for transanal total mesorectal excision. *Surg Endosc* 36(2):1143–1151
9. Padoy N (2019) Machine and deep learning for workflow recognition during surgery. *Minim Invasiv Ther* 28(2):82–90
10. Chen YW, Zhang J, Wang P, Hu ZY, Zhong KH (2022) Convolutional-de-convolutional neural networks for recognition of surgical workflow. *Front Comput Neurosci* 16:998096
11. García-Peraza-Herrera LC, Li W, Gruijthuisen C, Devreker A, Attilakos G, Deprest J et al (2020) Real-time segmentation of non-rigid surgical tools based on deep learning and tracking. Preprint at <https://arxiv.org/abs/2009.03016>

12. Jha D, Ali S, Tomar NK, Riegler MA, Johansen D, Johansen HD et al (2021) Exploring deep learning methods for real-time surgical instrument segmentation in laparoscopy. In: 2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), IEEE, Athens, pp 1–4
13. Twinanda AP, Shehata S, Mutter D, Marescaux J, de Mathelin M, Padoy N (2016) EndoNet: a deep architecture for recognition tasks on laparoscopic videos. *IEEE Trans Med Imaging* 36(1):86–97
14. Wang J, Jin Y, Cai S, Xu H, Heng PA, Qin J et al (2021) Real-time landmark detection for precise endoscopic submucosal dissection via shape-aware relation network. Preprint at <https://arxiv.org/abs/2111.04733>
15. Alexandrova S, Tatlock Z, Cakmak M (2015) RoboFlow: a flow-based visual programming language for mobile manipulation tasks. In: 2015 IEEE International Conference on Robotics and Automation (ICRA), IEEE, Seattle, pp 5537–5544
16. Allen B, Nistor V, Dutson E, Carman G, Lewis C, Faloutsos P (2009) Support vector machines improve the accuracy of evaluation for the performance of laparoscopic training tasks. *Surg Endosc* 24(1):170
17. Terven J, Cordova-Esparza D (2023) A comprehensive review of YOLO: from YOLOv1 to YOLOv8 and beyond. Preprint at <https://arxiv.org/abs/2304.00501>
18. Hashimoto DA, Rosman G, Rus D, Meireles OR (2018) Artificial intelligence in surgery: promises and perils. *Ann Surg* 268(1):70–76

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.