



Research article

Development of deep learning framework for anatomical landmark detection and guided dissection line during laparoscopic cholecystectomy

Pruittikorn Smithmaitrie^a, Methasit Khaonualsri^a, Wannipa Sae-Lim^b,
Piyanun Wangkulangkul^c, Supakool Jearanai^c, Siripong Cheewatanakornkul^{c,*}

^a Department of Mechanical and Mechatronics Engineering, Faculty of Engineering, Prince of Songkla University, Thailand

^b Department of Computer Science, Faculty of Science, Prince of Songkla University, Thailand

^c Minimally Invasive Surgery Unit, Department of Surgery, Faculty of Medicine, Prince of Songkla University, Thailand

ARTICLE INFO

Keywords:

Deep learning
Convolutional neural network
Anatomical landmark detection
Guided dissection line
Laparoscopic cholecystectomy
Bile duct injury

ABSTRACT

Background: Bile duct injuries during laparoscopic cholecystectomy can arise from misinterpretation of biliary anatomy, leading to dissection in improper areas. The integration of a deep learning framework into laparoscopic procedures offers the potential for real-time anatomical landmark recognition, ensuring accurate dissection. The objective of this study is to develop a deep learning framework that can precisely identify anatomical landmarks, including Rouviere's sulcus and the liver base of segment IV, and provide a guided dissection line during laparoscopic cholecystectomy.

Methods: We retrospectively collected 40 laparoscopic cholecystectomy videos and extracted 80 images from each video to establish the dataset. Three surgeons annotated the bounding boxes of anatomical landmarks on a total of 3200 images. The YOLOv7 model was trained to detect Rouviere's sulcus and the liver base of segment IV as anatomical landmarks. Additionally, the guided dissection line was generated between these two landmarks by the proposed algorithm. To evaluate the performance of the detection model, mean average precision (mAP), precision, and recall were calculated. Furthermore, the accuracy of the guided dissection line was evaluated by three surgeons. The performance of the detection model was compared to the scaled-YOLOv4 and YOLOv5 models. Finally, the proposed framework was deployed in the operating room for real-time detection and visualization.

Results: The overall performance of the YOLOv7 model on validation set and testing set were 98.1 % and 91.3 %, respectively. Surgeons accepted the visualization of guide dissection line with a rate of 95.71 %. In the operating room, the well-trained model accurately identified the anatomical landmarks and generated the guided dissection line in real-time.

Conclusions: The proposed framework effectively identifies anatomical landmarks and generates a guided dissection line in real-time during laparoscopic cholecystectomy. This research underscores the potential of using deep learning models as computer-assisted tools in surgery, providing an assistant tool to accommodate with surgeons.

* Corresponding author. Minimally Invasive Surgery Unit, Department of Surgery, Prince of Songkla University, Songkla, Thailand.
E-mail address: siripong.c@psu.ac.th (S. Cheewatanakornkul).

1. Introduction

Laparoscopic cholecystectomy (LC) is a minimally invasive surgical procedure that has become the gold standard for gallbladder removal. This minimally invasive approach provides several benefits including reduced postoperative pain, shortened hospital stays, and faster recovery periods compared to traditional open surgery [1]. However, a significant concern associated with LC is the risk of bile duct injury (BDI) [2], which can impact on the quality of life and survival of patients. BDI leads to prolonged hospitalization, re-operations, additional investigation, and increased treatment and reimbursement costs. Moreover, undiagnosed BDIs or delays in their treatment may potentially lead to medical malpractice lawsuits, placing an additional financial strain on the healthcare system. Implementing a well-defined safe surgical framework can help prevent these occurrences [3].

Considering advancements in surgical techniques, the incidence of BDI remains relatively high compared to the open surgical era [4], with reported rates ranging from 0.32 % to 0.52 % [5]. The primary cause of BDI is misidentification of the common bile duct as the cystic duct, which leads surgeons to perform dissection in potentially unsafe areas. This was true for both novice and experienced surgeons. Several strategies have been proposed for reducing BDI, including the critical view of safety (CVS) approach [6]. However, achieving CVS is occasionally accountable for BDI. Furthermore, the R4U line [7] was introduced in 2021 to divide the surgical anatomy into safe and unsafe areas. This line is drawn from Rouviere's sulcus to the liver base of segment IV. Consequently, based on the concept of the R4U line, fixed anatomical landmarks can serve as reference positions, and the R4U line will be created when the fixed anatomical landmarks are identified. To assist surgeons in identifying anatomical landmarks, the need of precise assistant tools is required. Therefore, it is essential to develop a verified method that can assist surgeons in real time during LC. Fig. 1 illustrates an example of the R4U line, along with key anatomical structures, including the gallbladder, Rouviere's sulcus, liver base of segment IV, cystic duct, common bile duct, and duodenum.

Deep learning, a subset of artificial intelligence, has been applied extensively in recent years to enhance medical image analysis. This computational approach enables machines to learn and make decisions based on trained data [8], exhibiting substantial promise for a broad range of medical applications. Skin cancer classification, for instance, has seen significant improvements through deep learning algorithms, facilitating better identification of cancerous appearances and improving overall patient prognosis [9]. Similarly, stroke detection has benefited from this approach, enabling rapid and precise identification of stroke lesions in Magnetic Resonance (MR) images, thereby accelerating patient treatment and potentially improving outcomes [10]. Deep learning has also been applied for ultrasound (US) image classification, making the analysis of Carotid artery disease, and enhancing diagnostic accuracy [11]. In addition, it has proven advantageous for cardiovascular disease (CVD) research, particularly in the classification of intima-media thickness (IMT), a key biomarker for early CVD [12]. In a more advanced application, COVID-Net was proposed for COVID-19 detection. COVID-Net achieved a high degree of sensitivity in the detection of COVID-19 cases [13]. Through these various applications, deep learning has demonstrated the potential to augment diagnostic precision, clinical decision-making, and patient management.

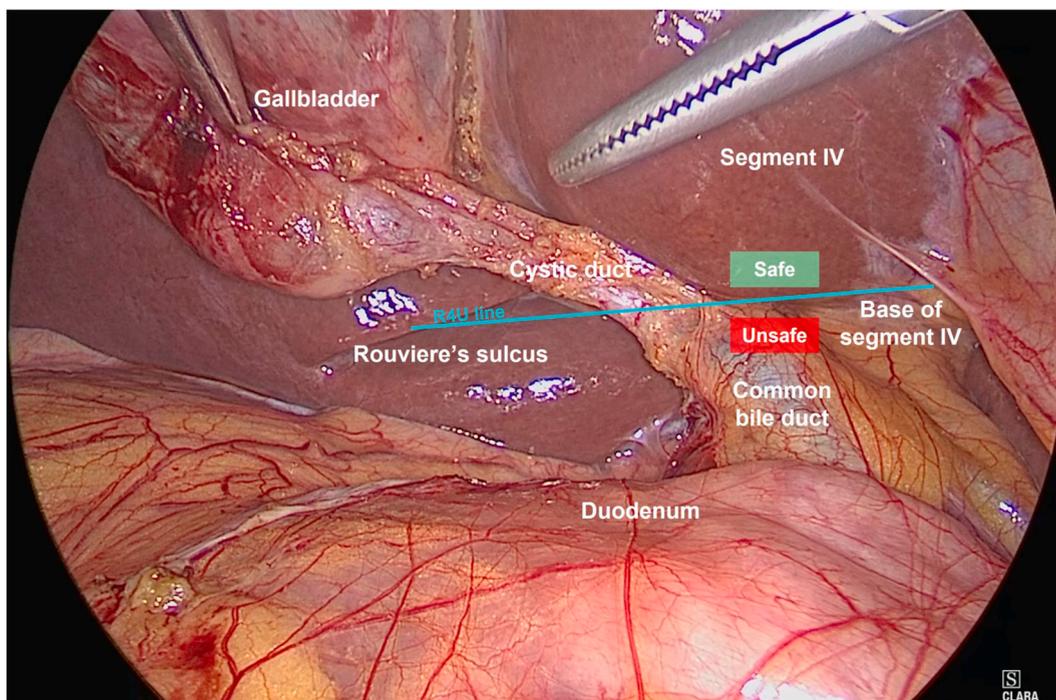


Fig. 1. The R4U line and key anatomical structures in a laparoscopic cholecystectomy image.

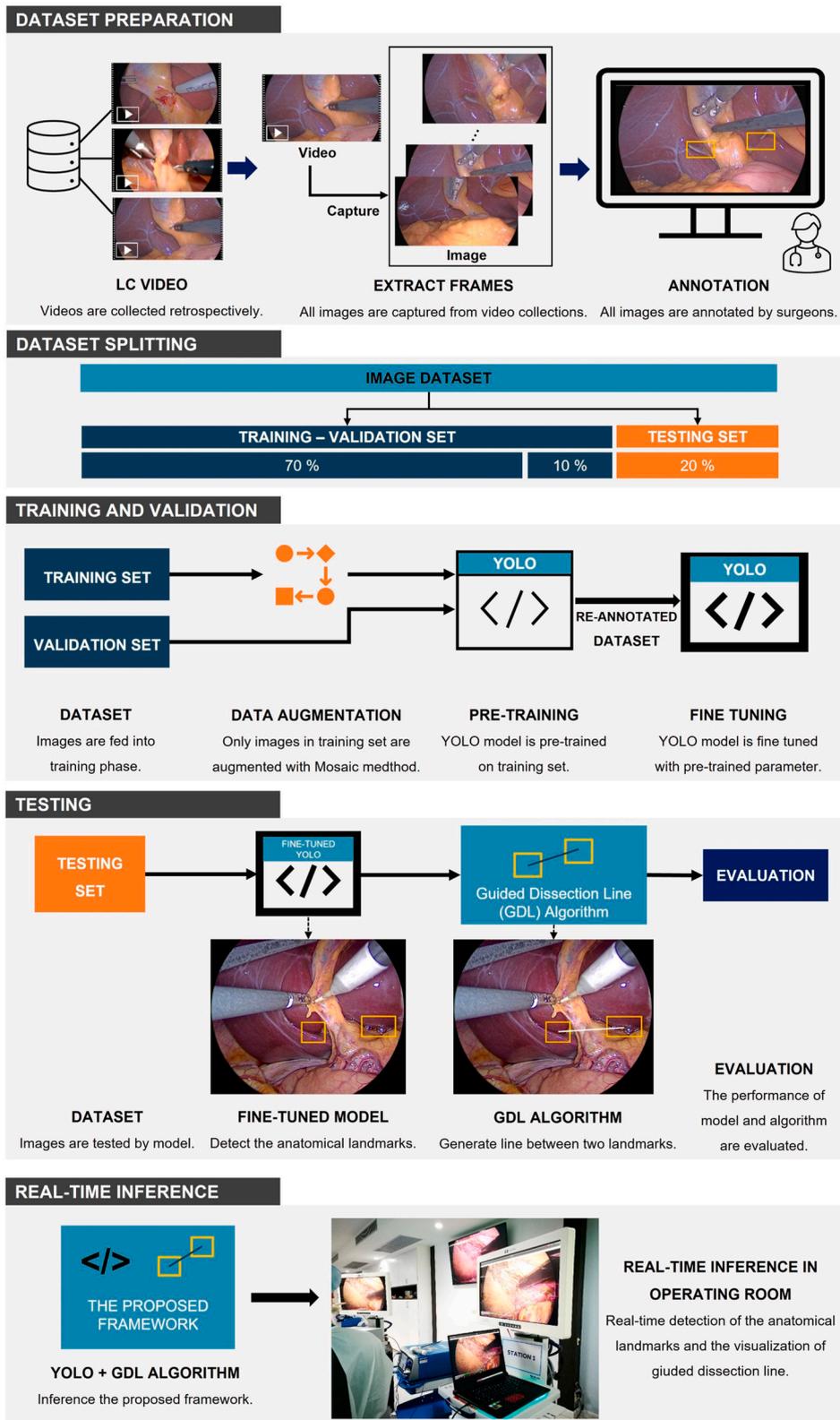


Fig. 2. Schematic representation of the proposed framework.

Incorporating these advancements into laparoscopic cholecystectomy, deep learning has the potential to assist surgeons in various aspects. The emerging field of surgical phase recognition using deep learning is evidenced in several studies [14–16], as is the identification of surgical instruments [17–19]. Moreover, numerous researchers have leveraged deep learning to detect anatomical landmarks during surgical procedures [20–22]. Tokuyasu et al. [23] devised a real-time object detection model based on YOLOv3 for identifying four landmarks, which displays the bounding box of the cystic duct, common bile duct, lower edge of the left medial segment, and Rouviere’s sulcus on monitors during LC. Additionally, GoNoGoNet [24], a convolutional neural network model, was introduced to identify safe and unsafe dissection zones by visualizing these areas with topographical heat maps. However, no previous research has specifically addressed the real-time visualization of a guided dissection line using deep learning techniques. To address this gap, this study introduces a novel approach by employing the R4U line concept to formulate a guide dissection line algorithm. This involves utilizing deep learning methodologies for the initial detection of anatomical landmarks.

The objective of this study is to propose a deep learning framework composed of a real-time detection model that is integrated with a guided dissection line algorithm. This framework is designed for the detection of two anatomical landmarks: Rouviere’s sulcus and the liver base of segment IV. Subsequent to the detection of these anatomical landmarks, the guided dissection line algorithm is activated to generate a visual guidance line between these two landmarks, delineating the safe zone for dissection during LC. To evaluate the proposed framework, we conducted a comprehensive analysis using a dataset from 40 LC cases. The underlying hypothesis of this study is that the deep learning framework could identify anatomical landmarks in real-time with high precision and could accurately visualize the guided dissection line. The study demonstrates that a deep learning model could be further applied to computer-assisted surgery, providing an assistant tool to collaborate with surgeons. This collaboration is crucial for enhancing surgical precision, reducing the risk of errors, and improving patient outcomes.

This paper is divided into five sections, as follows: Section 2 describes the materials and methods used in this study, including the dataset collection, the annotation process, the details of the deep learning model, the algorithm of the guide dissection line, and the process of model training, fine-tuning, and testing. Section 3 presents the experimental results demonstrating the performance and accuracy of the proposed framework. Section 4 provides a detailed discussion and analysis of this study. Finally, Section 5 summarizes the main contributions and highlights the importance of the research.

2. Materials and methods

The methodology of the proposed framework is illustrated in Fig. 2. The following subsections provide a detailed explanation of the methodology and its components.

2.1. Data collection and annotation

In this retrospective study, we acquired a collection of 40 laparoscopic cholecystectomy procedure videos from Songklanagarind Hospital, Prince of Songkla University. These procedures were conducted by a team of four surgeons, representing a broad spectrum of expertise, ranging from senior staff members to trainees. This dataset is characterized by its diversity, encompassing patients with a variety of demographic profiles, including different genders and age groups. Additionally, the video collection showcases a range of anatomical landmarks and utilizes endoscopic camera equipment from two manufacturers, Olympus and Karl Storz. This variety extends to the positioning and orientation of the cameras used during the surgeries.

A significant proportion of these cases were diagnosed as chronic cholecystitis based on histopathological analyses. The dataset also includes instances of subacute and acute cholecystitis, though it is noteworthy that no cases involving cancerous lesions were included. Adhering to strict ethical standards and privacy concerns, all personally identifiable information was removed and anonymized from the videos prior to their use in our research. This step was taken to ensure compliance with ethical guidelines and maintain the confidentiality of the patients involved.

In this study, our data engineering team captured 80 still images from each video of the gallbladder removal procedures. This

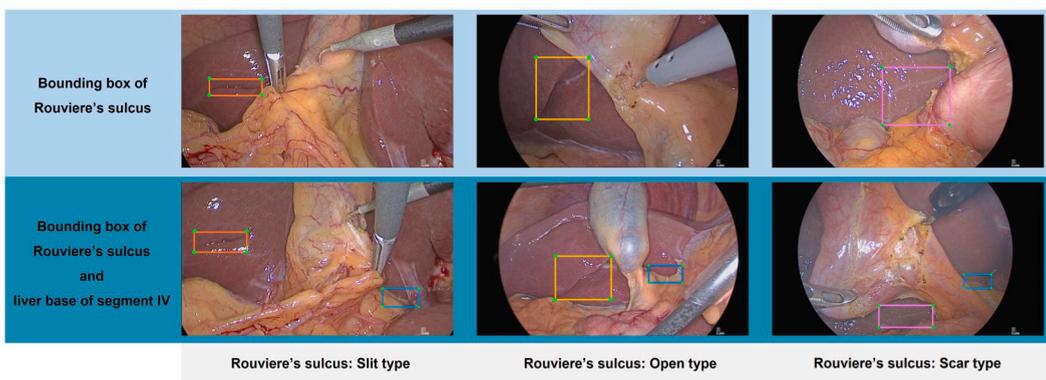


Fig. 3. Sample annotated anatomical landmarks in bounding boxes.

process was designed to encompass a comprehensive range of variabilities, such as differences in landmark positions, rotations, and scales. We also paid particular attention to capturing images with obstructed landmarks, aiming to create a robust and comprehensive dataset. Following the image capture phase, a team of three experienced surgeons undertook the task of annotating these images. A total of 3200 still images were annotated to delineate landmark bounding boxes and to classify object types.

The objects classified in this study were defined and included several anatomical landmarks: the liver base of segment IV, and three types of Rouviere’s sulcus - slit type, open type, and scar type. For the annotation process, we utilized LabelImg [25], a widely recognized open-source annotation tool. This tool enabled precise annotation of images, which were critical for the subsequent phases of our research. Each annotated image resulted in the creation of bounding boxes, the details of which were exported into .txt files. These files were specifically formatted to be compatible with the YOLO model, facilitating their direct application in training this machine learning model.

Fig. 3 shows a sample of the annotation for each anatomical landmark following the annotation description below:

- Liver base of segment IV: The bounding box is drawn to cover the lower edge of medial segment of the left lobe.
- Rouviere’s sulcus, slit type: The bounding box is drawn to cover the narrow and shallow sulcus.
- Rouviere’s sulcus, open type: The bounding box is drawn to cover the sulcus, which was seen to have its medial end open towards the hepatic portal.
- Rouviere’s sulcus, scar type: The bounding box is drawn to entirely cover the sulcus, which was obvious as a white scar.

The retrospective medical data collection and analysis was approved by Songklanagarind Hospital ethics committee with EC number REC.64-12-10-1.

2.2. Data splitting

From the total collection of 3200 images, the dataset was divided into three subsets: 2240 images for training, 320 for validation, and 640 for testing. A breakdown of the dataset reveals the following distribution of instances:

- Liver base of segment IV: Found in 1646 images.
- Rouviere’s sulcus, slit type: Present in 1166 images.
- Rouviere’s sulcus, open type: Captured in 1702 images.
- Rouviere’s sulcus, scar type: Depicted in 273 images.

Fig. 4 shows a visual representation of the distribution of each anatomical landmark across the training, validation, and testing sets.

2.3. Data pre-processing

Prior to training the detection model, all images in the dataset were resized from $1,920 \times 1,080$ pixels to 416×416 pixels utilizing the standard resizing function of the YOLOv7 model. Subsequently, the training set was enhanced by incorporating additional viewpoints of images through the Mosaic method [26]. During the Mosaic augmentation process, four distinct training images were chosen at random, partially cropped, and then combined into a single image to form augmented training batch data. This enables the

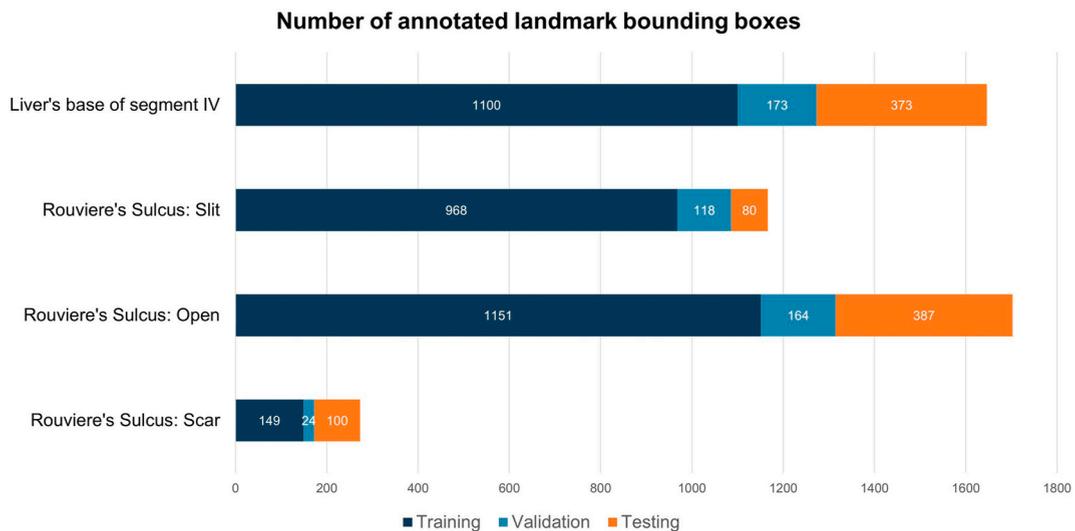


Fig. 4. Distribution of anatomical landmarks within the dataset.

model to acquire the objects at a lower scale than usual, reducing the requirement for a large batch size [27]. Furthermore, the risk of overfitting was reduced due to the diversity of perspectives in the augmented images [28]. Fig. 5 presents the batch results from the Mosaic augmentation method.

2.4. The YOLOv7 model

Alexey Bochkovskiy and Kin Yiu Wong recently released YOLOv7 [29], the outstanding real-time object detection model, to the YOLO family. The transformation of the YOLOv7 architecture achieved state-of-the-art accuracy and inference speed in comparison to previous versions of YOLO [30–33]. Fig. 6 illustrates the YOLOv7 architecture and details of each building block, as found in Ref. [34]. Typically, the YOLO architecture comprises three principal components: the backbone, neck, and head layer structures. Initially, the backbone extracts feature patches from the input images, which are then integrated at the neck layer before being passed to the head layer for prediction. The inference performance of model depends on the architecture of the backbone stage [35]. In the improvement of YOLOv7, an approach called Extended Efficient Layer Aggregation Networks (E-ELAN) has been introduced. E-ELAN enhances parameter utilization, computational blocks, and feature aggregation by adding the expand, shuffle, and merge cardinality process in the backbone stage. In addition, a novel compound scaling method has been proposed for concatenation-based models. This method scales the depth of a computation block and computes the width of the architecture by scaling the remaining transition layers in a specific ratio, thereby preserving the initial features and optimal structure. Lastly, the trainable bag-of-freebies concept, consisting of a re-parameterized convolution method and a coarse-to-fine lead guided assigner technique, has been introduced to potentially enhance model performance without increasing the training cost.

2.5. Training and fine-tuning of the model

In this research, the YOLOv7 model, utilized for training on the laparoscopic cholecystectomy dataset, comprises 415 layers with 37,212,738 parameters. The training process involved two phases: an initial training phase and a fine-tuning phase. During the initial training phase, the model was trained for 100 epochs without utilizing pre-trained weights. This allowed the model to learn from scratch specifically for the task of detecting anatomical landmarks in laparoscopic cholecystectomy. The predicted bounding box in this phase provided the characteristics that are useful for surgeons to re-annotate the training set. In the fine-tuning phase, the model was further trained for an additional 100 epochs using the re-annotated training set. We transferred the weights from the initial training phase to the fine-tuning phase, which is a common strategy in deep learning model training.

Learning rates were set at different values for the initial training and fine-tuning phases to establish a balance between learning speed and model stability. The learning rates during fine-tuning were reduced to prevent overfitting and allow for more precise adjustments. The batch size was determined considering the size of the dataset and the memory limitations of the GPU. Stochastic

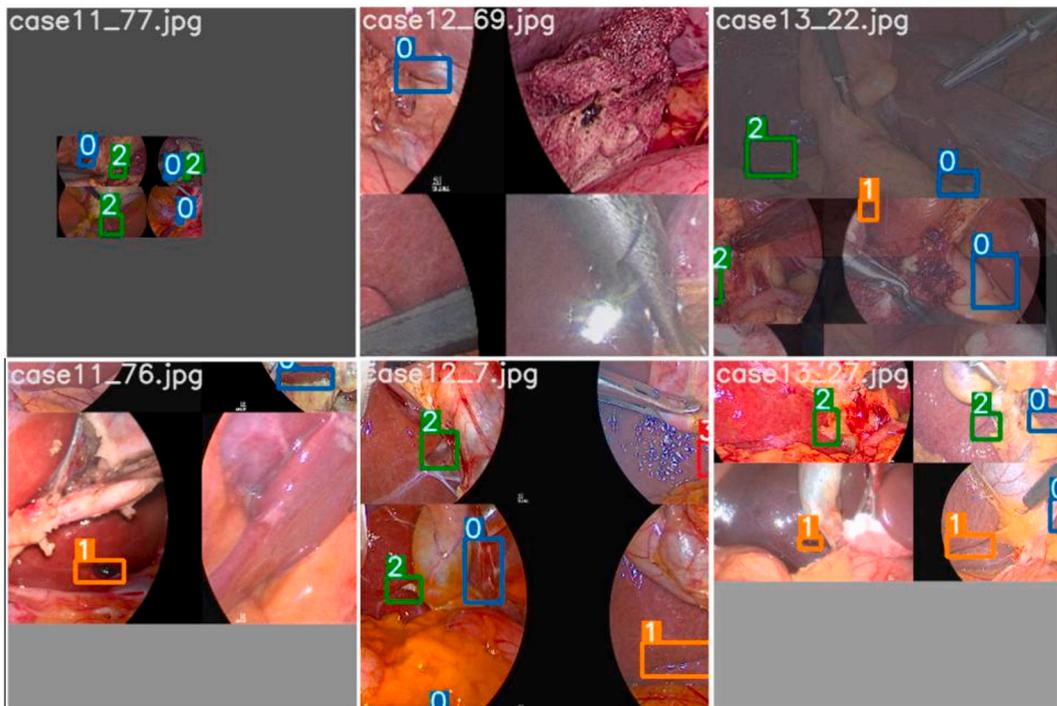


Fig. 5. Sample outputs from the Mosaic augmentation process.

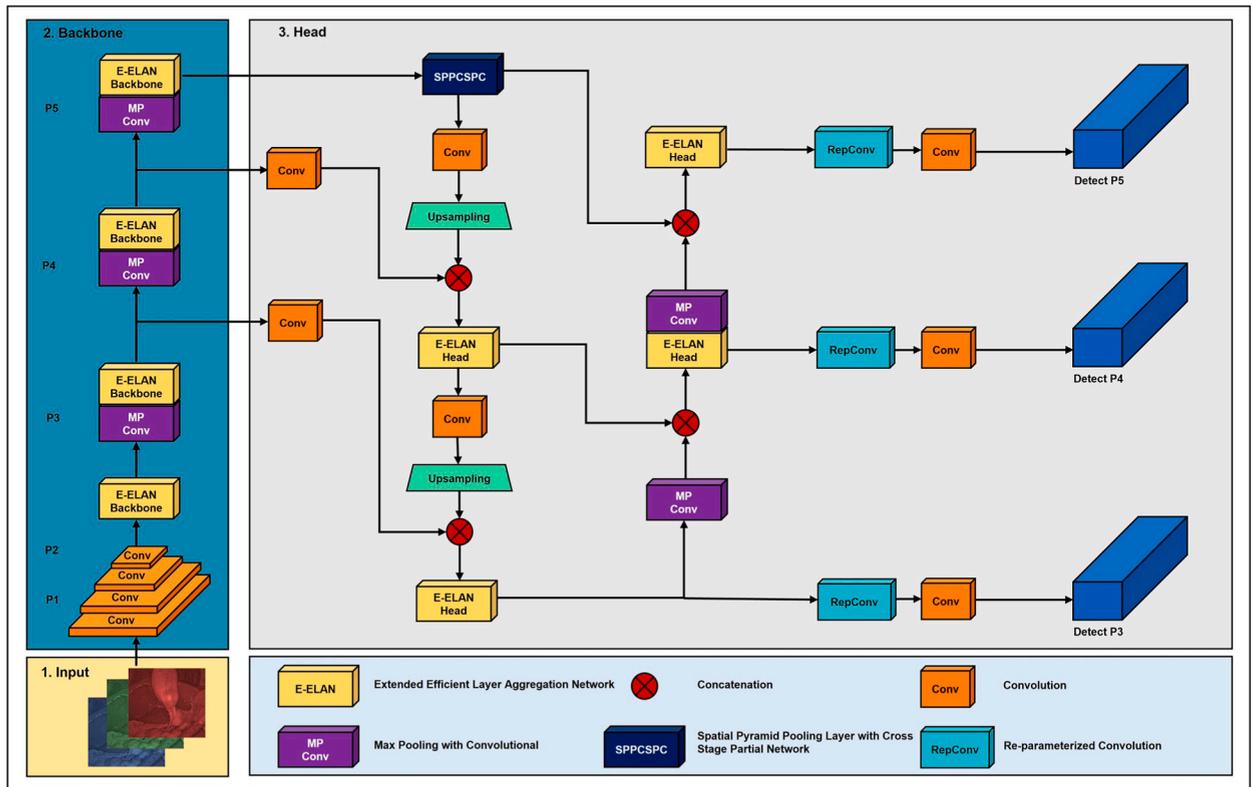


Fig. 6. The proposed framework architecture based on the YOLOv7 model.

Gradient Descent (SGD) was selected as the optimizer due to its efficiency in handling large datasets and its ability to avoid local minima during training. Based on our comparative analysis, SGD performs well in terms of convergence speed, optimization efficiency, and overall accuracy for the anatomical landmark detection task. Furthermore, momentum was defined to accelerate convergence and stabilize the training process, as it has been shown to be effective when used with SGD. To enhance the generalization of the model and mitigate the risk of overfitting, the Mosaic data augmentation technique was applied to generate a variety of training images.

The detailed hyperparameter settings, including the learning rates, batch size, optimizer, and data augmentation technique, are provided in Table 1. These settings were carefully defined to ensure effective model training and fine-tuning, ultimately leading to accurate detection of anatomical landmarks.

2.6. Guided dissection line algorithm

The guided dissection line algorithm was developed by applying the R4U line strategy [7] to generate a visual line between the predicted bounding boxes of the liver base of segment IV and Rouviere’s sulcus. This algorithm is activated when the predicted bounding boxes belong to different classes, specifically the liver base of segment IV and Rouviere’s sulcus. The algorithm draws a straight line from the center of one box to the center of the other, providing a reference line for surgeons to determine the safe dissection zone. Algorithm 1 outlines the implementation of the guided dissection line algorithm, and Fig. 7 provides a visual

Table 1
Detailed hyperparameters employed during the training process.

Hyperparameter	1st Training	2nd Training (Fine tuning)
Epoch	100	100
Batch size	16	16
Image size	416 × 416	416 × 416
Pre-trained weight	No	Transfer from 1st training
Initial learning rate	0.01	0.001
Final learning rate	0.1	0.1
Momentum	0.937	0.937
Optimizer	Stochastic Gradient Descent	Stochastic Gradient Descent
Data augmentation	Mosaic	Mosaic

representation of this algorithm.

Algorithm 1 Guided dissection line calculation and visualization

Input:

1. The coordinate of Rouviere's sulcus bounding box
(x_{RSmin} , y_{RSmin} , x_{RSmax} , y_{RSmax})
2. The coordinate of the liver base of segment IV bounding box
($x_{BsegIVmin}$, $y_{BsegIVmin}$, $x_{BsegIVmax}$, $y_{BsegIVmax}$)
3. The label of predicted class, $class[i]$

Output:

- The straight line between the center-point of two input boxes
- 1: **while** predicted class of bounding box $class[i] \neq class[i + 1]$ **do**
 - 2: compute the coordinate center-point of Rouviere's sulcus bounding box, C_{RS}
 - 3: $C_{RS} = \left[\left(\frac{x_{RSmin} + x_{RSmax}}{2}, \left(\frac{y_{RSmin} + y_{RSmax}}{2} \right) \right]$
 - 4: compute the coordinate center-point of liver base of segment IV bounding box, C_{BsegIV}
 - 5: $C_{BsegIV} = \left[\left(\frac{x_{BsegIVmin} + x_{BsegIVmax}}{2}, \left(\frac{y_{RSmin} + y_{RSmax}}{2} \right) \right]$
 - 6: Plot line between C_{RS} and C_{BsegIV}
 - 7: **end while**
-

2.7. Model testing and evaluation

Upon completion of the training phase, the model was subjected to testing using unseen images from the gallbladder removal procedure. The performance of the model for a particular class was evaluated using the precision, recall, and average precision (AP) metrics [36]. Meanwhile, the overall performance of the model was denoted by mean average precision (mAP), which represents the mean of the APs for all classes. The first step in calculating the mAP value is to compute intersection over union (IoU), as described in Equation (1), where B_{pred} is the predicted bounding box and B_{gt} is the ground-truth bounding box.

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

The IoU metric refers to the ratio of the intersection area to the union area of the predicted bounding box and the ground-truth bounding box, as illustrated in Fig. 8. Subsequently, the thresholds are defined to calculate the values of true positives (TP), false positives (FP), and false negatives (FN) by comparing the IoU with the defined thresholds. Then, the precision and recall values, along with their corresponding confidence levels, are calculated using Equations (2) and (3) for each class.

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

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

Next, the average precision (AP), the single value that is utilized to summarize the collection of precision and recall on the different

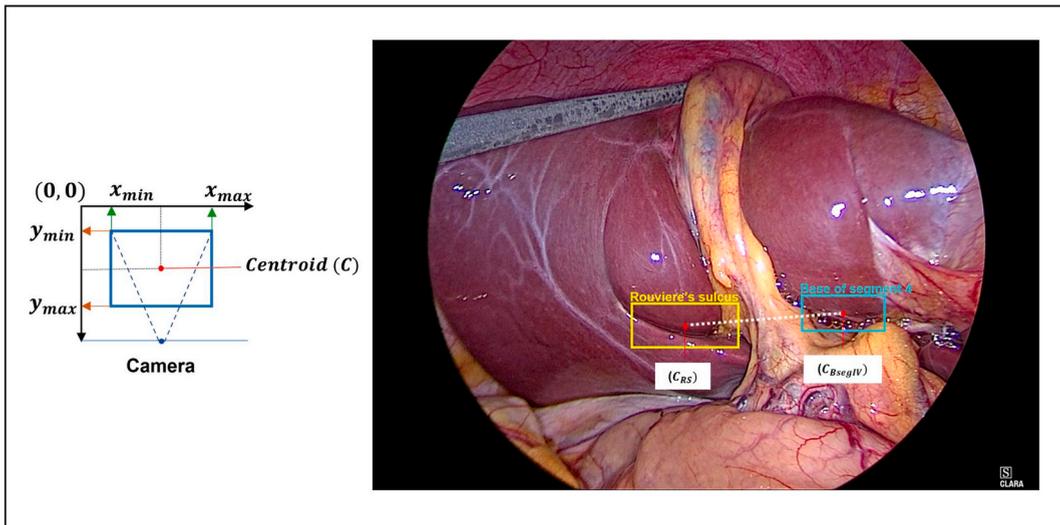


Fig. 7. Depiction of the guided dissection line algorithm.

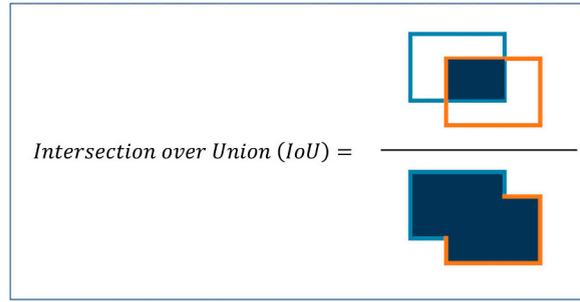


Fig. 8. Visualization of the IoU metric.

thresholds, is calculated according to Equation (4), where n is the number of threshold points, i is the index of each threshold, P_i is the collection of precision values, and R_i is the collection of recall values. Finally, the mAP is obtained by taking the mean of the AP of all classes, as shown in Equation (5), where AP_k is the AP of class k , and n is the number of classes.

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

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

3. Results

In this section, we present the experimental results and performance of the anatomical landmark detection using the YOLOv7 model, as well as the guided dissection line visualization based on the proposed algorithm. We compare the performance of YOLOv7 with its previous versions, scaled-YOLOv4 [32] and YOLOv5 [33]. Lastly, we explain the real-time inference results in the operating room.

3.1. Performance of anatomical landmark detection

The model for anatomical landmark detection underwent training for 100 epochs using the laparoscopic cholecystectomy dataset, followed by fine-tuning for another 100 epochs on the re-annotated dataset. The performance of the model for each anatomical landmark was assessed using precision, recall, and average precision (AP), while the mean average precision (mAP) was used to measure the overall performance of the model. These metrics were calculated using Equations (2)–(5), respectively.

The overall detection performance of the fine-tuned model on the validation set achieved a precision of 95.9 %, recall of 96.6 %, and mAP of 98.1 %. Similarly, on the testing set, the model achieved a precision of 91.7 %, recall of 86.5 %, and mAP of 91.3 %. Table 2 and Table 3 present the precision, recall, and AP of each landmark in the validation and testing sets, respectively. These tables demonstrate the performance of the detection model in identifying the liver base of segment IV and the different subclasses of Rouviere's sulcus. While the model achieved high precision and recall for all types of Rouviere's sulcus, the lower performance in detecting the liver base of segment IV indicates the difficulty in accurately identifying this anatomical area due to the similarity in pixel values between the liver base of segment IV and the area under Rouviere's sulcus, which is adjacent to the duodenum. The annotation of the bounding box for the liver base of segment IV contains pixels that resemble other areas. In contrast, the bounding box annotation for Rouviere's sulcus includes more distinct pixel values, enabling accurate prediction by the model. For a visual representation of the detection results, Fig. 9 presents the detection of anatomical landmarks on the testing set.

3.2. Performance of the guided dissection line

The guided dissection line visualization is enabled between two landmarks when the predicted bounding box appears at both the

Table 2

Evaluation of the anatomical landmark detection model across all classes on the validation set, with an IoU threshold of 0.5.

Class	Precision	Recall	AP
Liver base of segment IV	89.7 %	90.8 %	94.6 %
Rouviere's sulcus: Slit	98.7 %	97.5 %	98.9 %
Rouviere's sulcus: Open type	99.7 %	98.2 %	99.5 %
Rouviere's sulcus: Scar	95.7 %	100.0 %	99.4 %
Average	95.9 %	96.6 %	98.1 % (mAP)

Table 3
Evaluation of the anatomical landmark detection model across all classes on the testing set, with an IoU threshold of 0.5.

Class	Precision	Recall	AP
Liver base of segment IV	82.2 %	77.3 %	85.1 %
Rouviere’s sulcus: Slit	95.5 %	89.0 %	93.9 %
Rouviere’s sulcus: Open type	98.0 %	88.0 %	93.7 %
Rouviere’s sulcus: Scar	91.3 %	91.7 %	92.5 %
Average	91.7 %	86.5 %	91.3 % (mAP)

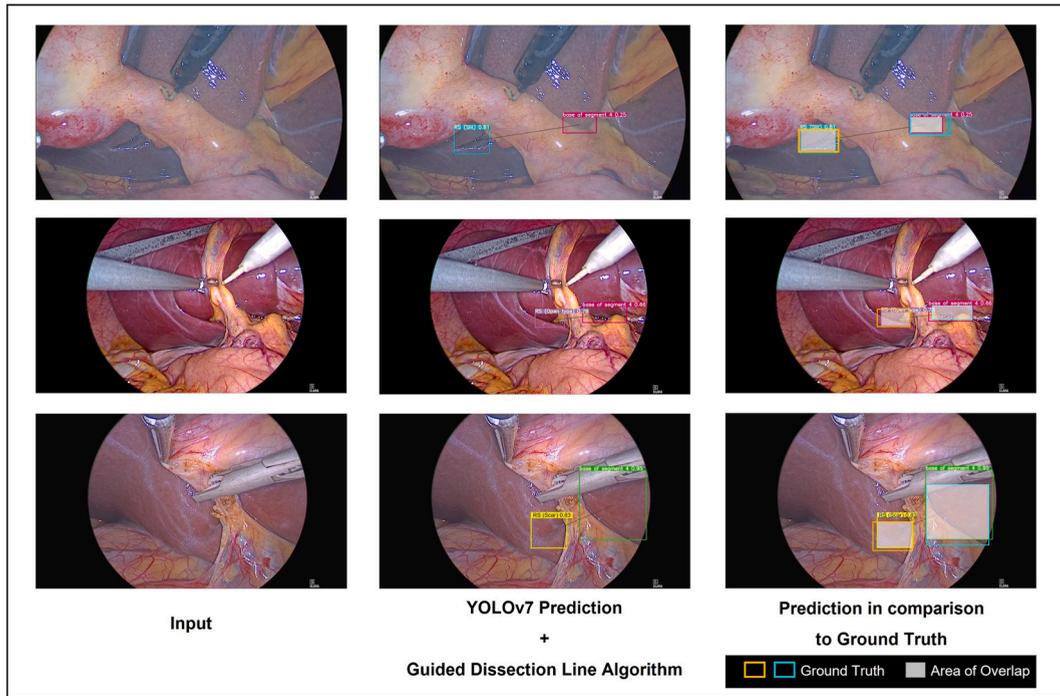


Fig. 9. Examples of input images (left), the detected anatomical landmarks and the guided dissection line visualization (middle), and the ground truth of anatomical landmarks (right).

liver base of segment IV and Rouviere’s sulcus landmark, as illustrated in Fig. 9. The outcome of the detection model directly influences the construction of the guided dissection line between these two landmarks. In this section, we evaluate the performance of the guided dissection line algorithm using the predicted results from the testing set. Out of the 640 images in the testing set, 360 annotated images contain both landmarks. However, the trained model correctly predicted both landmarks in 342 out of these 360 images. To evaluate the guided dissection line, these 342 images were assessed by three surgeons. Each surgeon provided a verdict, either accepting or rejecting the dissection line results based on safety concerns during surgery. Acceptance from a surgeon indicates agreement with the position of the guided dissection line. Conversely, rejection signifies dissatisfaction with its placement, primarily attributed to incorrectness in the visualization of the guided dissection lines. This misalignment, caused by improper anatomical landmarks prediction. The number of acceptances for the guided dissection line visualizations by each surgeon is presented in Table 4.

Table 4
Surgeons’ acceptances of the guided dissection line visualizations on the testing set.

	Number of acceptances	
	Guided dissection line visualizations on the testing set (342 images)	
Surgeon ID	Accept	Reject
Surgeon I	327	15
Surgeon II	336	6
Surgeon III	319	23
Average	327.33 (95.71 %)	14.66 (4.29 %)

3.3. Real-time application in the operating room

The framework proposed in this study was effectively implemented during laparoscopic cholecystectomy procedures at Songklanagarind Hospital. Live endoscopic video feeds, in a 4 K resolution (3840×2160 pixels), were sourced from the camera control unit (CCU) and relayed to a laptop. This setup facilitated the real-time detection of anatomical landmarks and the visualization of the guided dissection line. The real-time processing was executed on a laptop equipped with an Intel Core i7 processor, 16 GB RAM, and an NVIDIA GeForce 3080 graphics card. Fig. 10 shows the workstation setup within the operating room.

In the real-time inference phase, the trained model was capable of concurrently identifying the Rouviere's sulcus and the liver base of segment IV. It captured the latest images from live laparoscopic streams, which were continuously collected into a buffer system. These buffered images were then predicted at an average rate of 30.39 frames per second (FPS), achieving a time latency of approximately 32.90 ms per frame. This rapid processing capability was crucial for providing the real-time positioning of the practical guided dissection line during surgical procedures.

To implement our proposed framework in the operating room, we obtained approval for the clinical study protocol from the Human Research Ethics Committee, Faculty of Medicine, Prince of Songkla University, with REC no. 64-125-10-1. The protocol titled "Development of Artificial Intelligence for Prevention of Bile Duct Injury: Image analysis of Rouviere's sulcus and liver base of segment IV using machine learning" is in full compliance with International Guidelines for human research subject protection, such as the Declaration of Helsinki, Belmont Report, CIOMS Guideline, and the International Conference on Harmonization in Good Clinical Practice (ICH-GCP). The approval was granted on March 31, 2021.

4. Discussion

This research presents a real-time deep learning framework for the identification of anatomical landmarks and the visualization of a guided dissection line in LC. The detection model was trained and tested on a dataset of 40 LC cases, which were collected from LC videos and annotated by three experienced surgeons. The concept of the framework was developed based on the guidelines for a safe LC [37]. The Rouviere's sulcus and liver base of segment IV, a common landmark of hepatobiliary anatomy, can be used as a reference point to identify the safe area before beginning dissection [38,39], whereas the R4U line was adopted to divide the region into safe and unsafe dissection zones. Moreover, the use of deep learning for computer vision in surgery shows a successful opportunity of further development of a real-time navigation tool for surgeons during LC.

In terms of anatomical landmark detection, the YOLOv7 model is utilized to predict bounding box for Rouviere's sulcus and the liver base of segment IV. Table 5 presents a comparison of the detection performance metrics on the test set across three different models: Scaled-YOLOv4, YOLOv5, and YOLOv7. The comparison includes precision, recall, mAP at an IoU threshold of 0.5, and inference time for each of these models. The Scaled-YOLOv4 model achieved a precision of 83.1 %, a recall rate of 82.9 %, an mAP of 86.2 %, and had an inference time of 0.236 s per image. The YOLOv5 model, on the other hand, had slightly lower precision and recall rates at 81.4 % and 77.8 % respectively, and an mAP of 82.3 %, and required slightly more time to infer at 0.194 s per image. The YOLOv7 model outperformed both these models across all metrics. It demonstrated an improved precision rate of 91.7 %, and a higher recall rate of 86.5 %. Furthermore, the mAP was significantly higher at 91.3 %, indicating superior object detection performance. In particular, the YOLOv7 model also had the shortest inference time at 0.179 s per image, making it the most time-efficient of the three models.

In addition, the guided dissection line is visualized when the predicted bounding box is present simultaneously at the Rouviere's sulcus and liver base of segment IV. The output of the anatomical landmark detection model directly influences the establishment of the guided dissection line between these two landmarks. The high acceptance rate from three surgeons demonstrates the value of the guided dissection line as a practical tool for assisting surgeons during LC procedures.

Further to the improvement of model accuracy, the study focused on the representative images that the model required. Consequently, data augmentation and data re-annotation were applied to help the model predict more accurately. The Data augmentation method provides several perspectives of images from which the model can learn the various features and can prevent the model from overfitting [40]. After the initial training phase, the dataset was re-annotated by surgeons to ensure that it was properly annotated before continuing to the fine-tuning phase. This strategy can improve the consistency of data annotation and reduce the number of ambiguous bounding boxes that can result in false positive predictions [41].

The proposed framework was successfully deployed in the operation room. The key advantage of this framework is its ability to accurately present anatomical landmarks and guided dissection lines in real-time, allowing for continuous monitoring of safe zones during LC. In the real-time inference phase, the model collects the latest images from real-time laparoscopic streaming into a buffer system. These buffered images are then processed at a rate of 30.39 FPS, with an impressive time latency of approximately 32.90 ms per frame. This latency is imperceptibly short to the human eye, ensuring seamless real-time tracking and visualization, which is crucial during laparoscopic procedures. While the preliminary framework achieves promising outcomes, there are some limitations in terms of the imbalance class of Rouviere's sulcus type in the dataset. Typically, an open type of sulcus is the most frequently observed, followed by the slit type. In comparison, the scar type is rarely found among patients [42]. The dataset from the retrospective cases contains a few numbers of scar type appearance. Thus, to solve this problem, it will be required to not only apply the data augmentation method but also increase the number of examples featuring Rouviere's sulcus scar type in future studies.



Fig. 10. Workstation setup within the operating room.

Table 5

Performance comparison on the testing set (640 images).

Model	Precision	Recall	mAP (IoU = 0.5)	Inference time (Second per image)
Scaled-YOLOv4	83.1 %	82.9 %	86.2 %	0.236 s
YOLOv5	81.4 %	77.8 %	82.3 %	0.194 s
YOLOv7	91.7 %	86.5 %	91.3 %	0.179 s

5. Conclusions

This research proposed a deep learning-based framework that utilized the YOLOv7 model and a guide dissection line algorithm for real-time detection and surgical guidance during laparoscopic cholecystectomy. The integration of the YOLOv7 model and the guide dissection line algorithm provides surgeons with real-time detection of anatomical landmarks and a visual line for safe dissection. Furthermore, the comparative analysis has demonstrated that the YOLOv7 model outperforms scaled-YOLOv4 and YOLOv5 models in terms of accuracy and inference time. The YOLOv7 model has shown improved precision, recall, and mean average precision (mAP), confirming its significance as the primary detection model in the proposed framework. The acceptance rate of the guide dissection line visualization of the proposed framework was evaluated by three experienced surgeons. The results reveal a high average acceptance rate, indicating that surgeons considered the visualization to be precise and helpful during the procedure. In addition, real-time inference in the operating room has been successfully deployed with the anatomical landmark detection model and the guided dissection line algorithm. This was a significant achievement in establishing the line of safety. The proposed framework can act as a recommendation tool that can guide the surgeons to dissect in the appropriate areas.

Data availability statement

Dataset inquiry from this research will be made available on request.

CRediT authorship contribution statement

Pruittikorn Smithmaitrie: Conceptualization, Supervision, Writing – original draft, Writing – review & editing. **Methasit Khaonualsri:** Data curation, Formal analysis, Methodology, Software, Writing – original draft. **Wannipa Sae-Lim:** Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Piyanun Wangkulangkul:** Conceptualization, Data curation, Validation. **Supakool Jearanai:** Conceptualization, Data curation, Validation. **Siripong Cheewatanakornkul:** Conceptualization, Data curation, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by National Science, Research and Innovation Fund (NSRF) and Prince of Songkla University (Grant No. ENG6601312S). Authors would like to thank the surgical team in the operating room at Songklanagarind Hospital for their help and support in setting up the project station and their valuable comments throughout the operation.

References

- [1] D.E.M. Litwin, M.A. Cahan, Laparoscopic cholecystectomy, *Surg. Clin. North Am.* 88 (2008) 1295–1313, <https://doi.org/10.1016/j.suc.2008.07.005>.
- [2] S.M. Strasberg, Avoidance of biliary injury during laparoscopic cholecystectomy, *J. Hepato Biliary Pancreat. Surg.* 9 (2002) 543–547, <https://doi.org/10.1007/s005340200071>.
- [3] V. Gupta, G. Jain, Safe laparoscopic cholecystectomy: adoption of universal culture of safety in cholecystectomy, *World J. Gastrointest. Surg.* 11 (2019) 62–84, <https://doi.org/10.4240/wjgs.v11.i2.62>.
- [4] E.A. Shaffer, Epidemiology of gallbladder stone disease, *Best Pract. Res. Clin. Gastroenterol.* 20 (2006) 981–996, <https://doi.org/10.1016/j.bpg.2006.05.004>.
- [5] P.H. Pucher, L.M. Brunt, N. Davies, A. Linsk, A. Munshi, H.A. Rodriguez, A. Fingerhut, R.D. Fanelli, H. Asbun, R. Aggarwal, S.S.C.T. Force, Outcome trends and safety measures after 30 years of laparoscopic cholecystectomy: a systematic review and pooled data analysis, *Surg. Endosc.* 32 (2018) 2175–2183, <https://doi.org/10.1007/s00464-017-5974-2>.
- [6] S.M. Strasberg, L.M. Brunt, Rationale and use of the critical view of safety in laparoscopic cholecystectomy, *J. Am. Coll. Surg.* 211 (2010) 132–138, <https://doi.org/10.1016/j.jamcollsurg.2010.02.053>.
- [7] V. Gupta, G. Jain, The R4U planes for the zonal demarcation for safe laparoscopic cholecystectomy, *World J. Surg.* 45 (2021) 1096–1101, <https://doi.org/10.1007/s00268-020-05908-1>.
- [8] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A.W.M. van der Laak, B. van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88, <https://doi.org/10.1016/j.media.2017.07.005>.
- [9] D.K.A. AL-Saedi, S. Savaş, Classification of skin cancer with deep transfer learning method, *Comput. Sci.* (2022), <https://doi.org/10.53070/bbd.1172782>.
- [10] R.A.J. Alhatemi, S. Savaş, Transfer learning-based classification comparison of stroke, *Comput. Sci.* (2022), <https://doi.org/10.53070/bbd.1172807>.
- [11] S. Savaş, N. Topaloğlu, Ö. Kazıcı, P. Koşar, Comparison of deep learning models in Carotid artery intima-media thickness ultrasound images: CAIMTUSNet, *Bilişim Teknoloj. Derg.* 15 (2022) 1–12, <https://doi.org/10.17671/gazibtd.804617>.
- [12] S. Savaş, N. Topaloğlu, Ö. Kazıcı, P.N. Koşar, Performance comparison of Carotid artery intima media thickness classification by deep learning methods, *Int. Congr. Hum.-Comput. Interact., Optim. Robot. Appl. Proc.* (2019) 125–131, <https://doi.org/10.36287/setsci.4.5.025>.
- [13] A. Bhatt, A. Ganatra, K. Kotecha, COVID-19 pulmonary consolidations detection in chest X-ray using progressive resizing and transfer learning techniques, *Heliyon* 7 (2021) e07211, <https://doi.org/10.1016/j.heliyon.2021.e07211>.
- [14] D. Kitaguchi, N. Takeshita, H. Matsuzaki, H. Takano, Y. Owada, T. Enomoto, T. Oda, H. Miura, T. Yamanashi, M. Watanabe, D. Sato, Y. Sugomori, S. Hara, M. Ito, Real-time automatic surgical phase recognition in laparoscopic sigmoidectomy using the convolutional neural network-based deep learning approach, *Surg. Endosc.* 34 (2020) 4924–4931, <https://doi.org/10.1007/s00464-019-07281-0>.
- [15] A.P. Twinanda, S. Shehata, D. Mutter, J. Marescaux, M. de Mathelin, N. Padoy, EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos, 2016, <https://doi.org/10.48550/arxiv.1602.03012>. Arxiv.
- [16] K. Cheng, J. You, S. Wu, Z. Chen, Z. Zhou, J. Guan, B. Peng, X. Wang, Artificial intelligence-based automated laparoscopic cholecystectomy surgical phase recognition and analysis, *Surg. Endosc.* 36 (2022) 3160–3168, <https://doi.org/10.1007/s00464-021-08619-3>.
- [17] E. Colleoni, S. Moccia, X. Du, E.D. Momi, D. Stoyanov, Deep learning based robotic tool detection and articulation estimation with spatio-temporal layers, *Ieee Robotics Automation Lett* 4 (2019) 2714–2721, <https://doi.org/10.1109/ra.2019.2917163>.
- [18] Z. Zhao, S. Voros, Y. Weng, F. Chang, R. Li, Tracking-by-detection of surgical instruments in minimally invasive surgery via the convolutional neural network deep learning-based method, *Comput Assist Surg* 22 (2017) 26–35, <https://doi.org/10.1080/24699322.2017.1378777>.
- [19] B. Namazi, G. Sankaranarayanan, V. Devarajan, A contextual detector of surgical tools in laparoscopic videos using deep learning, *Surg. Endosc.* 36 (2022) 679–688, <https://doi.org/10.1007/s00464-021-08336-x>.
- [20] A. Madani, B. Namazi, M.S. Altieri, D.A. Hashimoto, A.M. Rivera, P.H. Pucher, A. Navarrete-Welton, G. Sankaranarayanan, L.M. Brunt, A. Okrainec, A. Alseidi, Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy, *Ann. Surg.* 276 (2020) 363–369, <https://doi.org/10.1097/sla.0000000000004594>.
- [21] A.A. Pozdeev, N.A. Obukhova, A.A. Motyko, Anatomical landmarks detection for laparoscopic surgery based on deep learning technology, 2021 Ieee Conf. Russ. Young Res. Electr. Electron. Eng. Elconrus (2021) 1668–1672, <https://doi.org/10.1109/elconrus51938.2021.9396093>, 00.
- [22] H. Nakanuma, Y. Endo, A. Fujinaga, M. Kawamura, T. Kawasaki, T. Masuda, T. Hirasaita, T. Etoh, K. Shinozuka, Y. Matsunobu, T. Kamiyama, M. Ishikake, K. Ebe, T. Tokuyasu, M. Inomata, An intraoperative artificial intelligence system identifying anatomical landmarks for laparoscopic cholecystectomy: a prospective clinical feasibility trial (J-SUMMIT-C-01), *Surg. Endosc.* (2022) 1–10, <https://doi.org/10.1007/s00464-022-09678-w>.
- [23] T. Tokuyasu, Y. Iwashita, Y. Matsunobu, T. Kamiyama, M. Ishikake, S. Sakaguchi, K. Ebe, K. Tada, Y. Endo, T. Etoh, M. Nakashima, M. Inomata, Development of an artificial intelligence system using deep learning to indicate anatomical landmarks during laparoscopic cholecystectomy, *Surg. Endosc.* 35 (2021) 1651–1658, <https://doi.org/10.1007/s00464-020-07548-x>.
- [24] S. Laplante, B. Namazi, P. Kiani, D.A. Hashimoto, A. Alseidi, M. Pasten, L.M. Brunt, S. Gill, B. Davis, M. Bloom, L. Pernar, A. Okrainec, A. Madani, Validation of an artificial intelligence platform for the guidance of safe laparoscopic cholecystectomy, *Surg. Endosc.* (2022) 1–9, <https://doi.org/10.1007/s00464-022-09439-9>.
- [25] Tzatalin LabelImg, Git Code, 2015. <https://github.com/heartexlabs/labelimg>.
- [26] W. Hao, S. Zhili, Improved mosaic: algorithms for more complex images, *J Phys Conf Ser* 1684 (2020) 012094, <https://doi.org/10.1088/1742-6596/1684/1/012094>.
- [27] F. Dadboud, V. Patel, V. Mehta, M. Bolic, I. Mantegh, Single-stage UAV detection and classification with YOLOV5: mosaic data augmentation and PANet, 2021 17th, Ieee Int Conf Adv Video Signal Based Surveillance Avss 00 (2021) 1–8, <https://doi.org/10.1109/avss52988.2021.9663841>.
- [28] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, *J Big Data* 6 (2019) 60, <https://doi.org/10.1186/s40537-019-0197-0>.
- [29] C.-Y. Wang, A. Bochkovskiy, H.-Y.M. Liao, YOLOv7: Trainable Bag-Of-Freebies Sets New State-Of-The-Art for Real-Time Object Detectors, 2022, <https://doi.org/10.48550/arxiv.2207.02696>. Arxiv.
- [30] J. Redmon, A. Farhadi, YOLOv3: an Incremental Improvement, 2018, <https://doi.org/10.48550/arxiv.1804.02767>. Arxiv.
- [31] A. Bochkovskiy, C.-Y. Wang, H.-Y.M. Liao, YOLOv4: Optimal Speed and Accuracy of Object Detection, 2020, <https://doi.org/10.48550/arxiv.2004.10934>. Arxiv.
- [32] C.-Y. Wang, A. Bochkovskiy, H.-Y.M. Liao, Scaled-YOLOv4: Scaling Cross Stage Partial Network, 2020, <https://doi.org/10.48550/arxiv.2011.08036>. Arxiv.
- [33] G. Joher, A. Chaurasia, A. Stoken, J. Borovec, NanoCode012, Y. Kwon, K. Michael, J. Fang TaoXie, Lorna Imyhyx, Z. Yifu, C. Wong, A. V. D. Montes, Z. Wang, C. Pati, J. Nadar, UnglvKitDe Laughing, V. Sonck, Yxnong Tkianai, P. Skalski, A. Hogan, D. Nair, M. Strobel, M. Jain, ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation, 2022, <https://doi.org/10.5281/zenodo.7347926>.
- [34] C.-Y. Wang, A. Bochkovskiy, H.-Y.M. Liao, YOLOv7: Trainable Bag-Of-Freebies Sets New State-Of-The-Art for Real-Time Object Detectors, 2022, <https://doi.org/10.48550/arXiv.2207.02696>. Arxiv.
- [35] N. Ma, X. Zhang, H.-T. Zheng, J. Sun, ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design, 2018, <https://doi.org/10.48550/arxiv.1807.11164>. Arxiv.

- [36] B. Özdemir, S. Aksoy, S. Eckert, M. Pesaresi, D. Ehrlich, Performance measures for object detection evaluation, *Pattern Recogn. Lett.* 31 (2010) 1128–1137, <https://doi.org/10.1016/j.patrec.2009.10.016>.
- [37] V. Gupta, G. Jain, Safe laparoscopic cholecystectomy: adoption of universal culture of safety in cholecystectomy, *World J. Gastrointest. Surg.* 11 (2019) 62–84, <https://doi.org/10.4240/wjgs.v11.i2.62>.
- [38] T.B. Hugh, New strategies to prevent laparoscopic bile duct injury—surgeons can learn from pilots, *Surgery* 132 (2002) 826–835, <https://doi.org/10.1067/msy.2002.127681>.
- [39] S. Connor, O.J. Garden, Bile duct injury in the era of laparoscopic cholecystectomy, *Br. J. Surg.* 93 (2006) 158–168, <https://doi.org/10.1002/bjs.5266>.
- [40] L. Perez, J. Wang, The Effectiveness of Data Augmentation in Image Classification Using Deep Learning, 2017, <https://doi.org/10.48550/arxiv.1712.04621>. Arxiv.
- [41] S.E. Whang, Y. Roh, H. Song, J.-G. Lee, Data Collection and Quality Challenges in Deep Learning: A Data-Centric AI Perspective, 2021, <https://doi.org/10.48550/arxiv.2112.06409>. Arxiv.
- [42] M. Singh, N. Prasad, The anatomy of Rouviere’s sulcus as seen during laparoscopic cholecystectomy: a proposed classification, *J. Minimal Access Surg.* 13 (2017) 89–95, <https://doi.org/10.4103/0972-9941.201731>.