

Original Article

Open Access



Real-time robust liver and gallbladder segmentation during laparoscopic cholecystectomy using convolutional neural networks: an analysis

Vahideh Ghobadi¹ , Luthffi Idzhar Ismail¹, Wan Zuha Wan Hasan¹, Haron Ahmad², Hafiz Rashidi Ramli¹, Nor Mohd Haziq Norsahperi¹, Anas Tharek³, Fazah Akhtar Hanapiah⁴

¹Faculty of Engineering, Universiti Putra Malaysia, Serdang 43400, Malaysia.

²KPJ Specialist Hospital Damansara, Petaling Jaya 47400, Malaysia.

³Hospital Sultan Abdul Aziz Shah, Universiti Putra Malaysia, Serdang 43400, Malaysia.

⁴Faculty of Medicine, Universiti Teknologi MARA (UiTM), Jalan Hospital, Sungai Buloh 47000, Malaysia.

Correspondence to: Dr. Luthffi Idzhar Ismail, Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, Jalan Universiti 2, Serdang 43400, Malaysia. E-mail: luthffi.ismail@upm.edu.my

How to cite this article: Ghobadi V, Ismail LI, Wan Hasan WZ, Ahmad H, Rashidi Ramli H, Norsahperi NMH, Tharek A, Hanapiah FA. Real-time robust liver and gallbladder segmentation during laparoscopic cholecystectomy using convolutional neural networks: an analysis. *Art Int Surg* 2024;4:278-87. <https://dx.doi.org/10.20517/ais.2024.30>

Received: 21 May 2024 **First Decision:** 4 Jul 2024 **Revised:** 13 Aug 2024 **Accepted:** 22 Aug 2024 **Published:** 8 Oct 2024

Academic Editors: Andrew A. Gumbs, Derek O'Reilly **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

Aim: Images in different laparoscopic cholecystectomy datasets are acquired using various camera models, parameters, and settings, with the annotation methods varying by institution. These factors result in inconsistent inference performance of the network model. This study aims to identify the optimal network model architecture for liver and gallbladder segmentation from several options. Then, the performance and robustness of the optimal network model are evaluated using an independent dataset that is not included in the training.

Methods: The public dataset, CholecSeg8k, was utilized as the input for the network model training, validation, and testing. A local private dataset from KPJ Damansara Hospital, Selangor, Malaysia, was used for testing purposes only. For the implementation of liver and gallbladder segmentation, segmentation models, a public Python library was employed.

Results: Among the experiments, highly accurate liver and gallbladder segmentation results were achieved using the feature pyramid network (FPN) architecture as the network model, with the Inception-ResNet-v2 architecture as the network backbone. The best-trained network model resulted in a loss of 0.070955, a mean intersection over



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



union (IoU) score of 0.95896, and a mean F1-score of 0.9773 on the test set. However, visualized results for the private dataset contained considerable false-negative areas.

Conclusion: The proposed automated technique has the potential to serve as an alternative to the conventional indocyanine green injection along with near-infrared fluorescence imaging (ICG-NIRF)-based method for liver and gallbladder segmentation during laparoscopic cholecystectomy. Future work will focus on enhancing the results of the private dataset. Additionally, a surgeon-assistant robotic arm that will use the liver and gallbladder segmentation results for camera steering will be analyzed.

Keywords: Artificial intelligence, convolutional neural network, deep learning, computer vision, liver and gallbladder segmentation, laparoscopic cholecystectomy

INTRODUCTION

During laparoscopic surgery, a laparoscopic camera serves as the surgeons' "eyes", enabling them to see the inside of the body and perform procedures without large incisions. Moreover, surgical navigation is a guideline for surgeons to make laparoscopic surgery safer and more efficient. Various navigation systems have been introduced to help surgeons visualize and navigate abdominal organs during laparoscopic surgery. Fu *et al.* investigated conventional non-artificial intelligence (AI)-based laparoscopic surgery navigation methods such as the use of computed tomography (CT)/magnetic resonance imaging (MRI) images taken before surgery, ultrasound images taken during surgery, indocyanine green injection along with near-infrared fluorescence imaging (ICG-NIRF) technology, and fusion of photos taken before and during surgery^[1]. However, each of these techniques may face its own set of challenges. In the ICG-based navigation method, Indocyanine Green, a liquid contrast, is injected into the patient's body. When ICG is irradiated by near-infrared light, it emits fluorescence radiation that can be detected and visualized by a special camera. The ICG-NIRF imaging method aids the surgeon to observe the surgical field by visualizing the hepatocystic structure effectively, as referenced by Wang *et al.* and Wendler *et al.*^[2,3]. Zaffino *et al.* discussed the drawbacks of the ICG-NIRF imaging technique^[4]. Spontaneous fluorescence radiation of healthy tissues may conflict with the fluorescence radiation of contrast agents in cancerous tissues. Therefore, it creates a low-quality video and image. An additional challenge is the time dedicated to imaging before or during laparoscopic surgery. The cost of this imaging technique is also considered an issue.

AI-based surgical navigation during laparoscopic surgery has emerged as a promising field. In the context of AI and computer vision, organ segmentation determines the boundary of an internal organ in medical images including X-rays, CT, MRI, ultrasound, and laparoscopic images. Accurate organ segmentation helps physicians and surgeons to detect diseases and make appropriate medical treatment decisions. An automated deep learning-based method could be an alternative to the conventional ICG-NIRF technique for segmenting the biliary structure during laparoscopic liver surgeries. The ICG-based method and symbolic representation of the deep learning-based approach are illustrated in [Figure 1](#). Such enhanced visualization can assist novice surgeons and aid them in training purposes. Additionally, it can be applied in the critical view of safety (CVS) assessment.

To help the surgeon perform a safe laparoscopic cholecystectomy, Madani *et al.* segmented the liver, gallbladder, and a few other parts of the hepatocystic structure in the frames were extracted from laparoscopic cholecystectomy videos^[5]. This segmentation was performed by applying the PSPNet network model designed for image segmentation. Scheikl *et al.* segmented five different classes, including liver and gallbladder, in laparoscopic cholecystectomy images^[6]. The performance of various network models was evaluated in their research. The best segmentation result was related to the ternausNet-11 network model

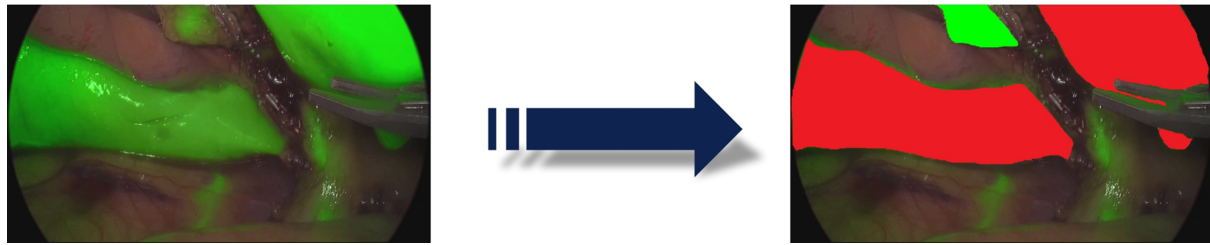


Figure 1. Liver segmentation by ICG-NIRF technique (left); Deep learning-based technique (right). The image on the right is created by photo editing software. ICG-NIRF: Indocyanine green injection along with near-infrared fluorescence imaging.

with the weights trained on the ImageNet dataset in the encoder part. Mascagni *et al.* performed segmentation for seven organs, including the gallbladder, in images captured from laparoscopic cholecystectomy videos using the DeeplabV3+ network model^[7]. Their research objective was to assess the CVS automatically for a safer laparoscopic cholecystectomy. The researchers have expanded their study. They have designed a software application called SurgFlow with AI-based capabilities to help surgeons during laparoscopic cholecystectomy^[8]. SurgFlow contains gallbladder segmentation in the subsection of hepatocystic anatomy segmentation by applying the DeepCVS network model. According to the surgeons' opinions engaged in this research, hepatocystic anatomy and surgical tool segmentation accompanied by phase detection were helpful for surgeons during laparoscopic cholecystectomy. It was due to a decrease in bile duct injuries caused by automated CVS assessment. The dataset utilized in their application was annotated following the annotation protocol offered by Mascagni *et al.*^[9]. Although, at the time of their research, SurgFlow was not practical in service by a surgeon during laparoscopic cholecystectomy, it might be considered a satisfactory attempt to advance AI in the laparoscopic cholecystectomy domain. Alt *et al.* proposed a technique to segment the gallbladder in 3D images captured by an RGB-D camera during robotic laparoscopic cholecystectomy^[10]. They performed gallbladder segmentation using the LapSeg3D network model (a modified 3D U-Net).

In AI-based applications, one of the important prerequisites is the availability of videos recorded during laparoscopic surgeries^[11]. In addition, labeled videos annotated by experts are required, while the annotation task is crucial in preparing surgical data^[12]. As a challenge, images of laparoscopic cholecystectomy datasets are captured by different camera parameters and settings. Moreover, annotation techniques vary among institutions. These limitations cause variations in the network model inference performance. Therefore, as presented in the literature^[8], there is a gap in the analysis of using different datasets for the training and testing stages to achieve robust organ segmentation during laparoscopic cholecystectomy.

The present study aims to analyze the prediction performance of several network models for real-time liver and gallbladder segmentation and to specifically assess the robustness of the best-trained network model identified through the experiments. The liver and gallbladder segmentation performance analysis was conducted by training various network models and backbones embedded in the open-source segmentation models Python library and utilizing different network parameters. This training was performed on the CholecSeg8k dataset, which consists of images extracted from laparoscopic cholecystectomy videos. To evaluate the robustness of the best-trained model, additional analysis was carried out using a private dataset provided by KPJ Damansara Hospital.

METHODS

In the present research, when the surgeon needs to identify the precise borders of the liver and gallbladder during laparoscopic cholecystectomy, the liver and gallbladder segmentation command is sent by the

surgeon by pressing a button or selecting an icon on a touchpad screen. The graphics processing unit (GPU) subsequently receives the image of the current surgical field of view from the laparoscopic camera. Then, this image is passed to the saved trained network model to predict liver and gallbladder segmentation in the image. Ultimately, the segmented liver and gallbladder are displayed on the surgeon's monitor. The overall flow of the network model training process (offline phase), along with the proposed automated real-time liver and gallbladder segmentation technique during laparoscopic cholecystectomy (online phase), is depicted in [Figure 2](#).

Dataset

For network model training, validation, and testing, a publicly online accessible dataset termed CholecSeg8k was utilized, along with a local private dataset from KPJ Damansara Hospital, Selangor, Malaysia. CholecSeg8k encompasses 8,080 laparoscopic cholecystectomy images extracted from 17 videos. All images are in portable network graphic (PNG) file format and sized at 854×480 . The annotations for these images include 13 classes, referencing Hong *et al.*^[13]. In this study, a subset of 4,040 images from the CholecSeg8k dataset were randomly selected. The data were divided into 2,828 images for training, 808 for validation, and 404 for testing, corresponding to 70%, 20%, and 10% of the selected images, respectively. The random image selection process and grouping were carried out by Python code. It is important to note that photos captured using the ICG-NIRI technique were excluded from the dataset. As part of the pre-processing steps, Python code was employed to resize and crop all images to 480×480 dimensions for the liver and gallbladder segmentation task. The private dataset was used just for the result testing.

Liver and gallbladder segmentation network model

Segmentation of more than one object in an image, such as the liver and gallbladder, is commonly referred to as multi-class segmentation but is termed segmentation for convenience here. Liver and gallbladder segmentation was implemented using the open-source segmentation models Python library, distributed under the MIT License^[14]. This library incorporates four distinct network models and 25 backbones. He *et al.* introduced a backbone as a pre-trained network model that provides pre-established weights to facilitate feature extraction during the encoding path^[15]. All network models and backbones available in segmentation models were trained, validated, and tested on the CholecSeg8k dataset. The best-trained network and backbone were feature pyramid network (FPN) and Inception-ResNet-v2, respectively. The FPN network model produces predictions at multiple scales. These predictions are combined to generate the final segmentation result. Therefore, the segmentation result obtained from FPN carries information that ranges (real-time) from fine to coarse details. The FPN network model is a suitable choice for object segmentation applications where objects vary in size^[16]. Inception-ResNet-v2 is a fusion of the Inception^[17] and ResNet^[18] network models in which Inception Residual blocks replace Inception blocks to prevent vanishing gradient obstacles in deeper network layers^[19]. The network loss function comprised the sum of dice loss and binary focal loss, with a scale equal to 1. Evaluation metrics included the intersection over union (IoU) score and the F1-score. The loss function, IoU score, and F1-score are defined as:

$$\text{loss} = \text{Dice Loss} + (1 \times \text{Binary Focal Loss}) \quad (1)$$

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|} \quad \text{where } A \text{ is the predicted bounding box and } B \text{ is the ground truth bounding box.} \quad (2)$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{where } \text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

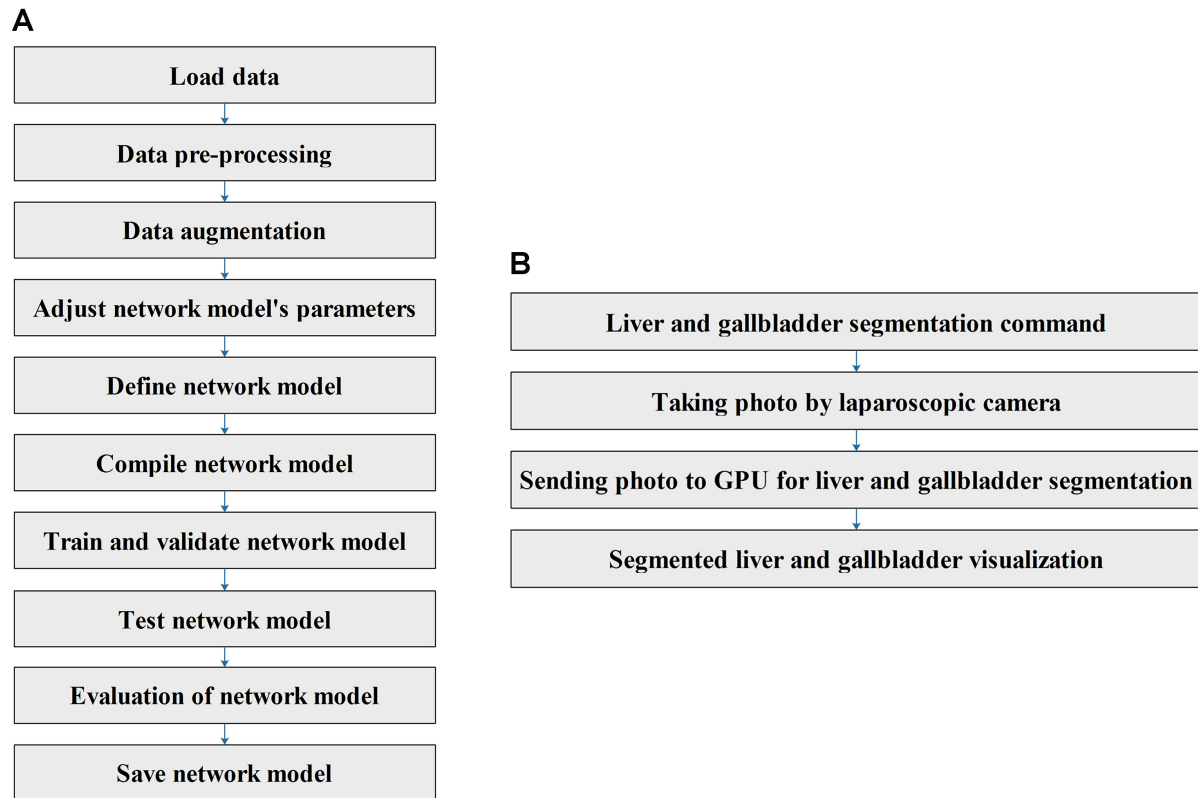


Figure 2. (A) Workflow of the network model training process (offline); (B) Workflow of automated liver and gallbladder segmentation.

In this research, the reported evaluation metrics and loss values reflected the average performance of liver and gallbladder segmentation calculated over the test set.

RESULTS

The training, validation, and test stages were conducted on an NVIDIA GeForce RTX 3060 GPU with 12 GB of RAM, utilizing Ubuntu 20.04 operating system and Python 3.9. Four network models with 25 backbones from the segmentation models Python library were employed in the training sessions. The number of epochs was considered to be 40 in different experiments. Observations indicated that the network overfits when the number of epochs exceeds 40. [Figure 3](#) reveals this point. The learning rate was adjusted to 0.01, 0.001, and 0.0001. Due to hardware constraints, batch sizes were consistently set to 4 in all experiments. All the backbones are pre-trained on the ImageNet dataset, leveraging existing weights. The loss, mean IoU score and mean F1-score express the average loss and evaluation metrics for liver and gallbladder segmentation, respectively. Their averages were computed over the entire test set. The liver and gallbladder segmentation experiments yielded the top three results from the three best-trained network models. The optimal liver and gallbladder segmentation outcome on the CholecSeg8k dataset was achieved using FPN as the network model, combined with Inception-ResNet-v2 as the backbone, with 40 epochs and a learning rate of 0.0001. Two more top results were achieved by keeping the number of epochs and learning rate the same as the top result but using a modified network model and backbone. [Table 1](#) details the segmentation results of the three best-trained network models. The presented loss and metrics are averages across the test set.

Table 1. Top three liver and gallbladder segmentation results from the top three best-trained network models

No.	Network model	Network backbone	Loss	Mean IoU score	Mean F1-score
1	FPN	Inception-ResNet-v2	0.070955	0.95896	0.9773
2	FPN	DenseNet201	0.0696	0.9583	0.97707
3	U-Net	EfficientNetB4	0.069792	0.95832	0.97692

IoU: Intersection over union; FPN: feature pyramid network.

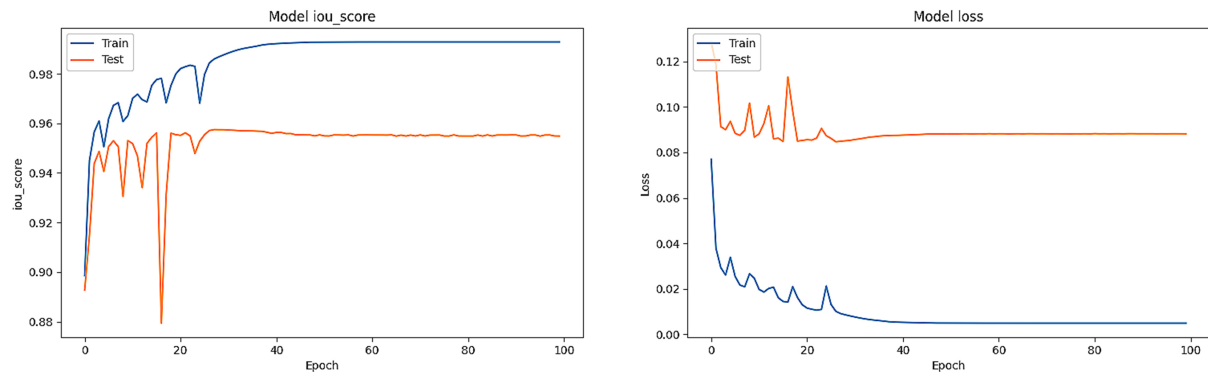


Figure 3. Plot of iou_score values in terms of epochs for training and validation (left); Plot of loss values in terms of epochs for training and validation (right).

To simulate real-time liver and gallbladder segmentation prediction, a webcam captured photos of color-printed cholecystectomy images. These images were selected from the CholecSeg8k dataset and were unseen by the network model. Afterward, photos were passed to the saved network model for segmentation prediction. Figure 4 visualizes the top three liver and gallbladder segmentation prediction results from the three best-trained network models. In Figure 5, the main image is overlaid with the liver and gallbladder segmentation results generated by the first top network model. In the continuation of this research, images from the private dataset were utilized in a separate experiment to validate the effectiveness and reliability of the best-trained network model for liver and gallbladder segmentation prediction. However, the result was not satisfactory. A random image from the private dataset was input into the saved first best-trained network model, and its liver and gallbladder segmentation prediction is illustrated in Figure 6. Due to the unavailability of labeled images for the private dataset, calculating the loss and metrics was impractical.

DISCUSSION

The present study investigates liver and gallbladder segmentation in laparoscopic cholecystectomy images using the segmentation models Python library. The FPN network model incorporating Inception-ResNet-v2 as its backbone outperformed other combinations. The F1-score and IoU score for the leading result were 0.9773 and 0.95896, respectively. It demonstrated that the segmentation models library generated promising liver and gallbladder segmentation results in laparoscopic images. The success could be attributed to the low data imbalance between the liver and non-liver areas in laparoscopic images of the CholecSeg8k dataset. The best-trained network model did not generate acceptable liver and gallbladder segmentation results on the private dataset and faced significant challenges. The challenges and their proposed solutions for segmentation improvement to accomplish a robust network model are discussed below.

- Some details of the images of the private dataset were lost when resized to the size of images in CholecSeg8k. Using a more effective image resizing algorithm will reduce information loss in resized images.

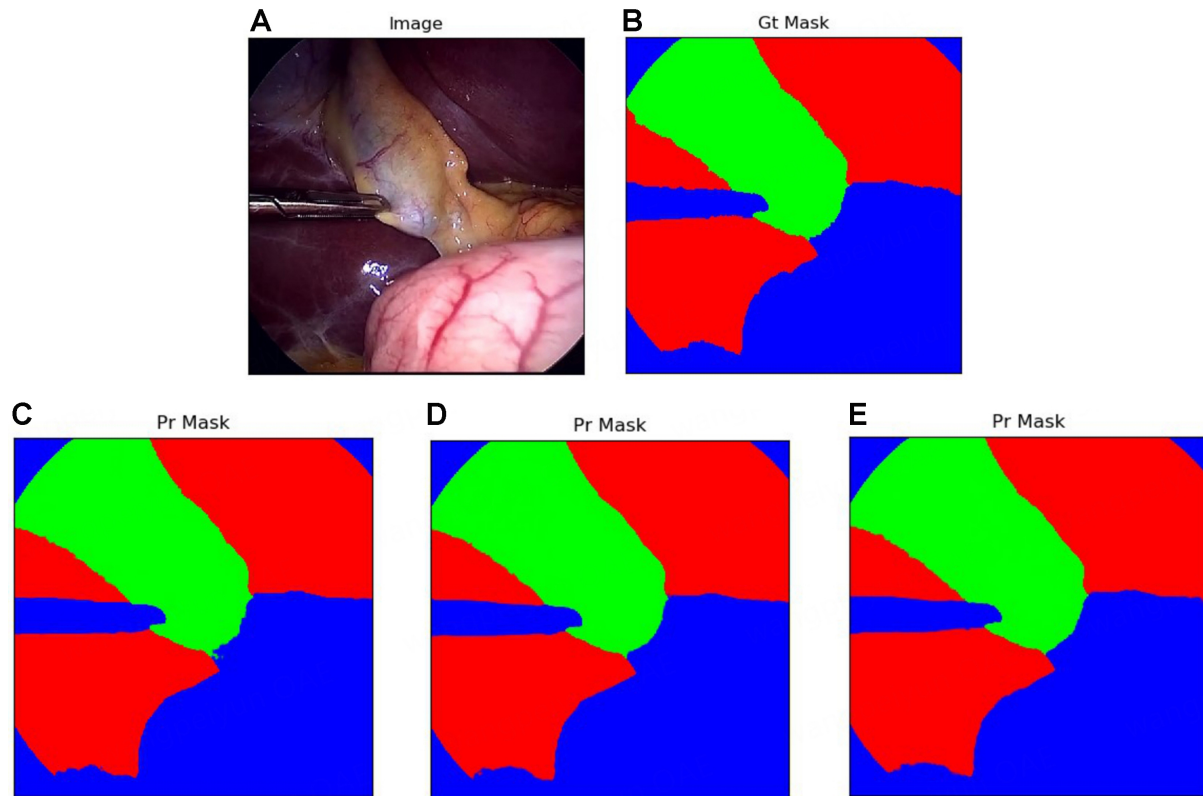


Figure 4. Liver and gallbladder segmentation results. (A) Original image; (B) Ground truth; (C) First top result; (D) Second top result; (E) Third top result. Areas with red, green, and blue colors depict the liver, gallbladder, and other classes (background), respectively.

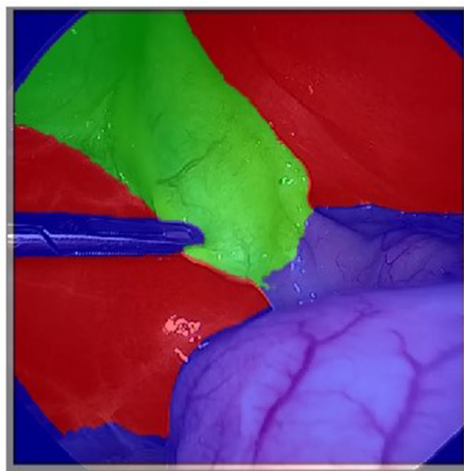


Figure 5. The main image overlaid by the first top liver and gallbladder segmentation results.

- In some parts, areas covered with blood were confused with the liver. This problem may be prevented by including blood segmentation in the network model training. It should also be considered that blood on the liver and gallbladder is not a separate organ and needs to be detected as a part of its container organ.

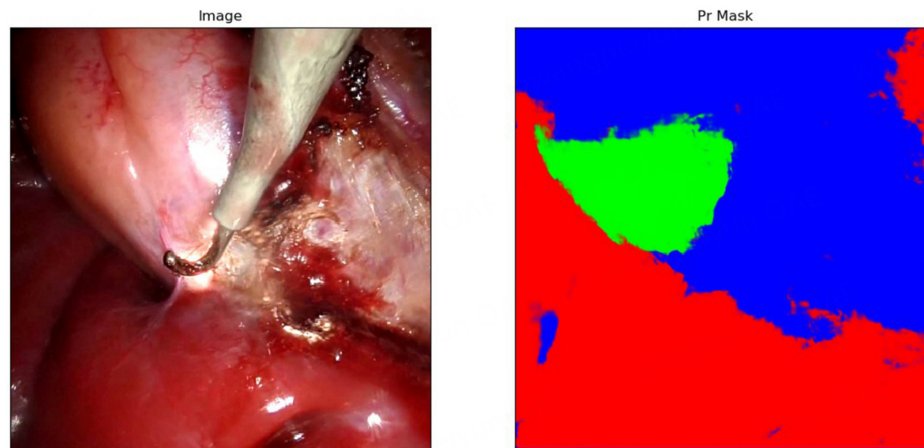


Figure 6. Random image from the private dataset (left); Segmented liver and gallbladder (right). The red, green, and blue colors correspond to the liver, gallbladder, and other classes (background), respectively.

Table 2. Comparison of the network model details for liver and gallbladder segmentation between Madani *et al.* and current research^[5]

Ref.	Dataset	Number of frames	Network model	Network backbone	Mean IoU score	Mean F1-score
Madani <i>et al.</i> ^[5]	Videos from 136 institutions	2,627	PSPNet ^[20]	ResNet50	0.79	0.88
Current	CholecSeg8k	4,040	FPN	Inception-ResNet-v2	0.95896	0.9773

IoU: Intersection over union; FPN: feature pyramid network.

- White areas caused by lighting should be considered part of the containing organ.
- In the case that images from other datasets are used just in the network model evaluation, domain adaption techniques can be employed to refine the results.
- Due to the unavailability of annotated images for the private dataset, the fine-tuning method is not applicable.

A comparative analysis was conducted between Madani *et al.* and the top result of the current research^[5]. The mentioned article was one of the few papers with available liver and gallbladder segmentation results. The details of this comparison are elaborated in Table 2. The mean IoU score and mean F1-score represent the average liver and gallbladder segmentation performance metrics on the test set. The IoU score and mean F1-score in the first row were calculated by averaging the individual scores for liver and gallbladder segmentation reported by Madani *et al.*^[5].

The strength of this research is the combination of FPN with the Inception-ResNet-v2 backbone that produced acceptable liver and gallbladder segmentation results. Using a local private dataset for liver and gallbladder segmentation enhanced collaboration among AI experts, data scientists, and medical professionals in local hospitals. However, a limitation of this study was its reliance on a single dataset for training and validation, which undermined the generalization of the segmentation prediction and robustness of the network model.

Future research endeavors will refine the segmentation result through domain adaptation techniques to address challenges in using images from different datasets. The performance of the network model will be further validated through larger-scale clinical trials. Moreover, a robotic arm that leverages the liver and gallbladder segmentation results will be analyzed to assist surgeons in camera steering during laparoscopic cholecystectomy.

In conclusion, the proposed automated technique offers a promising alternative to the traditional ICG-NIRF-based method for improved visualization of the liver and gallbladder during laparoscopic cholecystectomy. This enhanced visualization will benefit novice surgeons, be valuable for educational purposes, and potentially aid in CVS assessment.

DECLARATIONS

Acknowledgments

Thanks to KPJ Damansara Specialist Hospital for collaborating and providing data.

Authors' contributions

Made substantial contributions to the conception and design of the study and performed data analysis and interpretation: Ghobadi V, Ismail LI

Performed data acquisition, as well as providing administrative, technical, and material support: Ahmad H Technical Support, and administrative support: Wan Hasan WZ

Administrative support: Rashidi Ramli H, Norsahperi NMH, Tharek A, Hanapiah FA

Availability of data and materials

The CholecSeg8k dataset, is accessible online. The local private dataset cannot be shared as it belongs to KPJ Damansara Hospital, Selangor, Malaysia.

Financial support and sponsorship

This work was supported by the Fundamental Research Grant Scheme, the Ministry of Higher Education (FRGS/1/2022/TK07/UPM/02/14.). The funding body had no role in the design of the experiment, the collection, analysis, and interpretation of data, or the writing of the manuscript.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2024.

REFERENCES

1. Fu Z, Jin Z, Zhang C, et al. The future of endoscopic navigation: a review of advanced endoscopic vision technology. *IEEE Access* 2021;9:41144-67. DOI
2. Wang X, Teh CSC, Ishizawa T, et al. Consensus guidelines for the use of fluorescence imaging in hepatobiliary surgery. *Ann Surg* 2021;274:97-106. DOI PubMed
3. Wendler T, van Leeuwen FWB, Navab N, van Oosterom MN. How molecular imaging will enable robotic precision surgery: the role

- of artificial intelligence, augmented reality, and navigation. *Eur J Nucl Med Mol Imaging* 2021;48:4201-24. DOI PubMed PMC
4. Zaffino P, Moccia S, De Momi E, Spadea MF. A review on advances in intra-operative imaging for surgery and therapy: imagining the operating room of the future. *Ann Biomed Eng* 2020;48:2171-91. DOI PubMed
 5. Madani A, Namazi B, Altieri MS, et al. Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. *Ann Surg* 2022;276:363-9. DOI PubMed PMC
 6. Scheikl PM, Laschewski S, Kisilenko A, et al. Deep learning for semantic segmentation of organs and tissues in laparoscopic surgery. *Curr Dir Biomed Eng* 2020;6:20200016. DOI
 7. Mascagni P, Vardazaryan A, Alapatt D, et al. Artificial intelligence for surgical safety: automatic assessment of the critical view of safety in laparoscopic cholecystectomy using deep learning. *Ann Surg* 2022;275:955-61. DOI PubMed
 8. Mascagni P, Alapatt D, Lapergola A, et al. Early-stage clinical evaluation of real-time artificial intelligence assistance for laparoscopic cholecystectomy. *Br J Surg* 2024;111:znad353. DOI PubMed
 9. Mascagni P, Alapatt D, Garcia A, et al. Surgical data science for safe cholecystectomy: a protocol for segmentation of hepatocystic anatomy and assessment of the critical view of safety. arXiv. [Preprint.] Sep 20, 2021. [accessed on 2024 Sep 21]. Available from: <https://doi.org/10.48550/arXiv.2106.10916>.
 10. Alt B, Kunz C, Katic D, et al. LapSeg3D: weakly supervised semantic segmentation of point clouds representing laparoscopic scenes. In: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2022 Oct 23-27; Kyoto, Japan. IEEE; 2022. pp. 5265-70. DOI
 11. Mascagni P, Alapatt D, Sestini L, et al. Computer vision in surgery: from potential to clinical value. *NPJ Digit Med* 2022;5:163. DOI PubMed PMC
 12. Maier-Hein L, Eisenmann M, Sarikaya D, et al. Surgical data science - from concepts toward clinical translation. *Med Image Anal* 2022;76:102306. DOI PubMed PMC
 13. Hong WY, Kao CL, Kuo YH, Wang JR, Chang WL, Shih CS. CholecSeg8k: a semantic segmentation dataset for laparoscopic cholecystectomy based on Cholec80. arXiv. [Preprint.] Dec 23, 2020. [accessed on 2024 Sep 21]. Available from: <https://doi.org/10.48550/arXiv.2012.12453>.
 14. Iakubovskii P. Segmentation_models. 2019. Available from: https://github.com/qubvel/segmentation_models. [Last accessed on 21 Sep 2024].
 15. He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. In: 2017 IEEE International Conference on Computer Vision (ICCV); 2017 Oct 22-29; Venice, Italy. IEEE; 2017. pp. 2980-8. DOI
 16. Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. arXiv. [Preprint.] Apr 19, 2017. [accessed on 2024 Sep 21]. Available from: <https://doi.org/10.48550/arXiv.1612.03144>.
 17. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27-30; Las Vegas, USA. IEEE; 2016. pp. 2818-26. DOI
 18. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27-30; Las Vegas, USA. IEEE; 2016. pp. 770-8. DOI
 19. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-ResNet and the impact of residual connections on learning. In: 31st AAAI Conference on Artificial Intelligence. 2017. pp. 4278-84. DOI
 20. Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017 Jul 21-26; Honolulu, USA. IEEE; 2017. pp. 6230-9. DOI