

# Deep learning-based localization of the biliary tract in laparoscopic images acquired during surgical robotic procedures

Cristina Iacono<sup>1</sup>, Fanny Ficuciello<sup>1</sup>, Umberto Bracale<sup>2</sup>

<sup>1</sup>*Department of Electrical Engineering and Information Technology*

<sup>2</sup>*Department of General Vascular and Thoracic Surgery*

*Università degli Studi di Napoli Federico II*

Naples, Italy

cristina.iacono@unina.it, fanny.ficuciello@unina.it, umbertobracale@gmail.com

Sara Moccia

*The BioRobotics Institute*

*Scuola Superiore Sant'Anna*

Pisa, Italy

sara.moccia@santannapisa.it

Aldo Marzullo

*Department of Mathematics and Computer Science*

*Università della Calabria*

Rende, Italy

marzullo@mat.unical.it

Elena De Momi

*Information and Bioengineering*

*Politecnico di Milano*

Milan, Italy

elena.demomi@polimi.it

**Abstract**—Laparoscopic cholecystectomy is a minimally invasive procedure whereby the gallbladder is removed using laparoscopic techniques. With more than 500,000 cholecystectomies performed per year, great interest has developed in laparoscopic cholecystectomy. The major advantages of laparoscopic cholecystectomy with respect to traditional cholecystectomies are the short hospital stay and early return to normal activity. Morbidity is low, but there is a concern about bile duct injuries. This study proposes a deep learning-based algorithm for localization of the biliary tract during laparoscopic surgical procedures. A dataset has been collected, consisting of videos of standard surgical practice, and has been used to train the deep learning algorithm.

**Index Terms**—deep learning, robotic surgery, localization

## I. INTRODUCTION

Cholecystectomy is one of the most frequently performed procedures in gastrointestinal surgery, and the laparoscopic approach is now the gold standard for symptomatic cholelithiasis and chronic and acute cholecystitis. Since the introduction of Laparoscopic cholecystectomy (LC), surgeons have prioritized preventing complications.

Besides the advantages of a distinctly faster recovery and better cosmetic results, the laparoscopic approach bears a higher risk for bile duct injury with an incidence in the range 0.3–1.5%. Bile duct injury has a significant impact on quality of life and survival. Many measures have been implemented to reduce the risk of bile duct injury during laparoscopic cholecystectomy.

To avoid bile duct injury, intraoperative visualization of the bile duct using near-infrared light and the fluorescent dye indocyanine green (ICG) were introduced during cholecystectomy [1]. ICG is metabolized by the liver and excreted in bile, making it an excellent medium for biliary imaging. The

problem in using ICG is that, while enhancing the bile duct, it makes it challenging to see all the other anatomical structures.

This work aims to address this problem, helping the surgeon better visualize the biliary tract without the use of ICG. To this end, a deep-learning algorithm for the localization of the biliary tract from white-light images acquired during the standard surgical practice has been implemented. This work also includes the construction and annotation of an image database to train the deep learning algorithm.

## II. METHODS

The method directly uses laparoscopic images to localize the biliary duct. To this end, You Only Look Once (YOLO), a state-of-the-art convolutional neural network, has been used. [2] YOLO is a regression-based object detector that looks at the whole image once to perform the detection. It consists of a single CNN that simultaneously predicts bounding boxes and the class probabilities for the boxes. In addition, YOLO looks at the entire image to encode contextual information during prediction, and thus, it is extremely fast and found suitable to detect or localize objects in real-time.

An image dataset has been collected from 12 video clips of 12 different patients who underwent laparoscopic cholecystectomy during 2020. The videos collected from patients that presented complications that did not fall within the scope of this study were rejected. The videos were acquired through a high-definition endoscopic camera system with a 25 Hz frame rate during surgical endoscopic procedures. The frames extracted from the videos were sampled one every ten frames, obtaining 399 frames. The frames were then manually annotated by drawing a bounding box on the bile duct. The video frames were split to have 208 frames for the training set

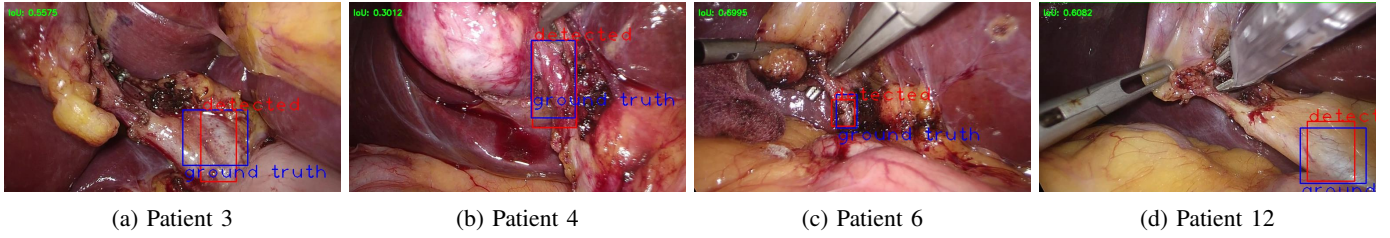


Fig. 1: Results of the localization algorithm on 4 different patients. Patient 3 (a), 6 (c) and 12 (d) belong only on the test set.

and 191 frames for the test set, as illustrated in Table I. The training set was used to train the neural network, while the test set was used for evaluation purposes only. The frames of three patients out of twelve (Patients 3, 6 and 12) have been used only in the test set.

TABLE I: Dataset composition

	<i>Total Frames</i>	<i>Training</i>	<i>Test</i>
Patient 1	142	15	15
Patient 2	171	34	14
Patient 3	219	-	39
Patient 4	152	74	20
Patient 5	48	5	5
Patient 6	144	-	29
Patient 7	168	14	10
Patient 8	89	18	10
Patient 9	153	14	10
Patient 10	73	20	10
Patient 11	27	14	10
Patient 12	135	-	19

For the training phase the following parameters were used:

- Batch size: 64,
- Learning rate: 0.001,
- Epochs number: 1748,
- Average loss: 0.18.

The Intersection over Union (IoU) was used as evaluation metrics: IoU compares the annotated bounding boxes with the bounding boxes predicted by the network, as follows:

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (1)$$

### III. RESULTS

To confirm the performance of the proposed method, the network has been used on the test set. Table II shows the results of the biliary duct detection on each patient. Based on the experimental results, the overall IoU is 0.67.

Regarding the videos, which have frames in the training set, the worst-case scenario happened for Patient 4, where the lowest values of the IoU are found. This is probably due to the fact that the ground truth bounding boxes have been annotated manually, and therefore in some cases, they may be less precise than others, and also because in the video, the biliary tract is not easily distinguishable from the background.

TABLE II: Detection Results

	<i>IoU</i>	<i>STD</i>
Patient 1	0.65	0.04
Patient 2	0.70	0.07
Patient 3	0.53	0.02
Patient 4	0.63	0.11
Patient 5	0.73	0.02
Patient 6	0.65	0.04
Patient 7	0.80	0.01
Patient 8	0.63	0.08
Patient 9	0.76	0.02
Patient 10	0.84	0.02
Patient 11	0.65	0.02
Patient 12	0.58	0.04

However, as can be seen from Fig. 1 (b), the localization of the surgical site of interest is accurate enough for the aim of the work.

The frames of patients 3, 6 and 12 were used only in the test set. The algorithm recognized the area of interest in 26 of 29 images in video 6 and in 14 of 19 images in video 12. The worst-case scenario happened in video 3 where only 6 of 39 images are correctly recognized.

### IV. CONCLUSION

This work addresses the problem of biliary tract injury during laparoscopic cholecystectomy, using an innovative approach in relation to the work suggested by the literature. The method proposes the application of YOLO for the localization of the biliary tract, creating a dataset of annotated frames of the surgical scene. The average IoU is equal to 67%, despite the small size of the dataset. The future goal is to use the localization of the biliary tract in real time and implement an augmented reality system, in order to help the surgeon to correctly and more easily recognize the area of interest in the crucial phases.

### REFERENCES

- [1] S. Vlek, D. Van Dam, S. Rubinstein, E. de Lange-de Klerk, L. Schoonmade, J. Tuynman, W. Meijerink, and M. Ankersmit, "Biliary tract visualization using near-infrared imaging with indocyanine green during laparoscopic cholecystectomy: results of a systematic review," *Surgical endoscopy*, vol. 31, no. 7, pp. 2731–2742, 2017.
- [2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.