



OPEN ACCESS

EDITED BY

Richard Jaepyeong Cha,
George Washington University, United States

REVIEWED BY

Fabrizio Consorti,
Sapienza University of Rome, Italy
Nir Horesh,
Sheba Medical Center, Israel

*CORRESPONDENCE

Xuehai Ding
✉ dinghai@shu.edu.cn

[†]These authors have contributed equally to this work

RECEIVED 13 January 2024

ACCEPTED 08 April 2024

PUBLISHED 19 April 2024

CITATION

Sang T, Yu F, Zhao J, Wu B, Ding X and Shen C (2024) A novel deep learning method to segment parathyroid glands on intraoperative videos of thyroid surgery.
Front. Surg. 11:1370017.
doi: 10.3389/fsurg.2024.1370017

COPYRIGHT

© 2024 Sang, Yu, Zhao, Wu, Ding and Shen. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A novel deep learning method to segment parathyroid glands on intraoperative videos of thyroid surgery

Tian Sang^{1†}, Fan Yu^{2†}, Junjuan Zhao¹, Bo Wu³, Xuehai Ding^{1*} and Chentian Shen²

¹School of Computer Engineering and Science, Shanghai University, Shanghai, China, ²Department of Nuclear Medicine, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China, ³Department of Thyroid, Breast and Hernia Surgery, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

Introduction: The utilization of artificial intelligence (AI) augments intraoperative safety and surgical training. The recognition of parathyroid glands (PGs) is difficult for inexperienced surgeons. The aim of this study was to find out whether deep learning could be used to auxiliary identification of PGs on intraoperative videos in patients undergoing thyroid surgery.

Methods: In this retrospective study, 50 patients undergoing thyroid surgery between 2021 and 2023 were randomly assigned (7:3 ratio) to a training cohort ($n = 35$) and a validation cohort ($n = 15$). The combined datasets included 98 videos with 9,944 annotated frames. An independent test cohort included 15 videos (1,500 frames) from an additional 15 patients. We developed a deep-learning model Video-Trans-U-HRNet to segment parathyroid glands in surgical videos, comparing it with three advanced medical AI methods on the internal validation cohort. Additionally, we assessed its performance against four surgeons (2 senior surgeons and 2 junior surgeons) on the independent test cohort, calculating precision and recall metrics for the model.

Results: Our model demonstrated superior performance compared to other AI models on the internal validation cohort. The DICE and accuracy achieved by our model were 0.760 and 74.7% respectively, surpassing Video-TransUnet (0.710, 70.1%), Video-SwinUnet (0.754, 73.6%), and TransUnet (0.705, 69.4%). For the external test, our method got 89.5% precision 77.3% recall and 70.8% accuracy. In the statistical analysis, our model demonstrated results comparable to those of senior surgeons (senior surgeon 1: $\chi^2 = 0.989$, $p = 0.320$; senior surgeon 2: $\chi^2 = 1.373$, $p = 0.241$) and outperformed 2 junior surgeons (junior surgeon 1: $\chi^2 = 3.889$, $p = 0.048$; junior surgeon 2: $\chi^2 = 4.763$, $p = 0.029$).

Discussion: We introduce an innovative intraoperative video method for identifying PGs, highlighting the potential advancements of AI in the surgical domain. The segmentation method employed for parathyroid glands in intraoperative videos offer surgeons supplementary guidance in locating real PGs. The method developed may have utility in facilitating training and decreasing the learning curve associated with the use of this technology.

KEYWORDS

artificial intelligence, deep learning, parathyroid glands, intraoperative videos, localization, segmentation

1 Introduction

The parathyroid gland (PG) is the smallest endocrine organ in the human body and plays an important role in maintaining the balance of calcium metabolism. Identifying and safeguarding PGs constitute a pivotal aspect of thyroid surgery. Given their small size and close anatomical proximity to lymph nodes or adipose tissues, there exists a risk of PG damage and compromised blood supply during thyroid surgery. Unfamiliarity with the morphology and anatomy of PGs heightens the susceptibility to damage, resulting in hypocalcemia that adversely affects the patient's quality of life. The injury of PGs can easily lead to postoperative hypocalcemia and other complications. One of the most common complications after total thyroidectomy is hypoparathyroidism. According to existing research, the incidence rate is about 30%–60% (1–3). It has been reported that the incidence of accidental parathyroid removal is up to 12%–28% (4). The symptoms of PG injury mainly include hand and foot twitching, limb sensory abnormalities, muscle spasms, and even life-threatening conditions.

Distinguishing PGs from similar tissues like thyroid, lymph nodes, or brown adipose tissues poses a significant challenge. Although all surgeons learn to differentiate PGs from other tissues, the length of this learning curve may be different for each surgeon, depending on experience. Even experienced surgeons cannot guarantee that all PGs will be recognized and protected in every thyroid surgery. At present, there are some methods to recognize PGs, such as carbon nanoparticles negative development, parathyroid hormone (PTH) in fine-needle aspiration (FNA) washout fluids, pathology validation, or near-infrared autofluorescence (NIRAF) (5–7). However, these methods have some limitations, such as longer periods of detection, high costs, false-positive results. Therefore, the existing recognition skills of PGs still rely on experienced surgeons. Nevertheless, early recognition of PGs before dissection is very helpful to guide dissection. In this case, a visual algorithm that recognizes the shapes and localization of a PG in the surgical field based on deep learning would be useful to shorten the learning curve and guide dissection.

To address those issues, we proposed the Video-Trans-U-HRNet model. At present, video AI technology has developed rapidly and is widely used in the medical field, such as robot surgery (8). Our task is to locate and segment the PG during thyroidectomy. To our knowledge, the application of AI methods to the localization and segmentation of PGs has mainly been limited to static images. However, thyroid surgery is a dynamic process that requires the full attention of the surgeon, and intraoperative videos can better reflect the real situation of the PG during surgery. Furthermore, the position and shape of PGs can undergo deformation and displacement due to the surgeon's manipulation, emphasizing the need for automatic localization and segmentation during the procedure. This not only enhances precision but also helps prevent accidental injury to PGs by the surgeon. In thyroid surgery, the PGs exhibit similarities to many other tissues, creating a challenge for existing AI models employed in medical segmentation and detection tasks to accurately recognize them. The objective of our study is to test a new AI-based method for recognizing PG and compare its performance to both junior and senior surgeons, with the ultimate aim of enhancing patient safety.

2 Methodology

2.1 Developing the AI model

The entire method is illustrated in Figure 1, comprising four parts: (A) The overall data flow of our method. (B) The detailed structure of our AI model, consisting of the Temporal Contextual Module (TCM), Encoder part, and Decoder part. (C) The dataset employed in this study. (D) The legends used in this figure. Zeng et al.'s work demonstrated the applicability of classical semantic segmentation models in medical segmentation. Drawing inspiration from their Video-TransUNet, we devised a deep-learning segmentation method tailored for the specific task of PGs video segmentation (9). The TransUnet model demonstrates robust performance in organ segmentation tasks, achieving outstanding results across diverse public datasets (10). Building upon TransUnet, we introduced a TCM to establish links between contextual information across frames for video segmentation (11). As depicted in Figure 1B, the TCM module comprises three parallel frames processing modules, all sharing a softmax layer. This module adeptly captures information from both preceding and succeeding frames, skillfully integrating it for feature extraction of the current key frame.

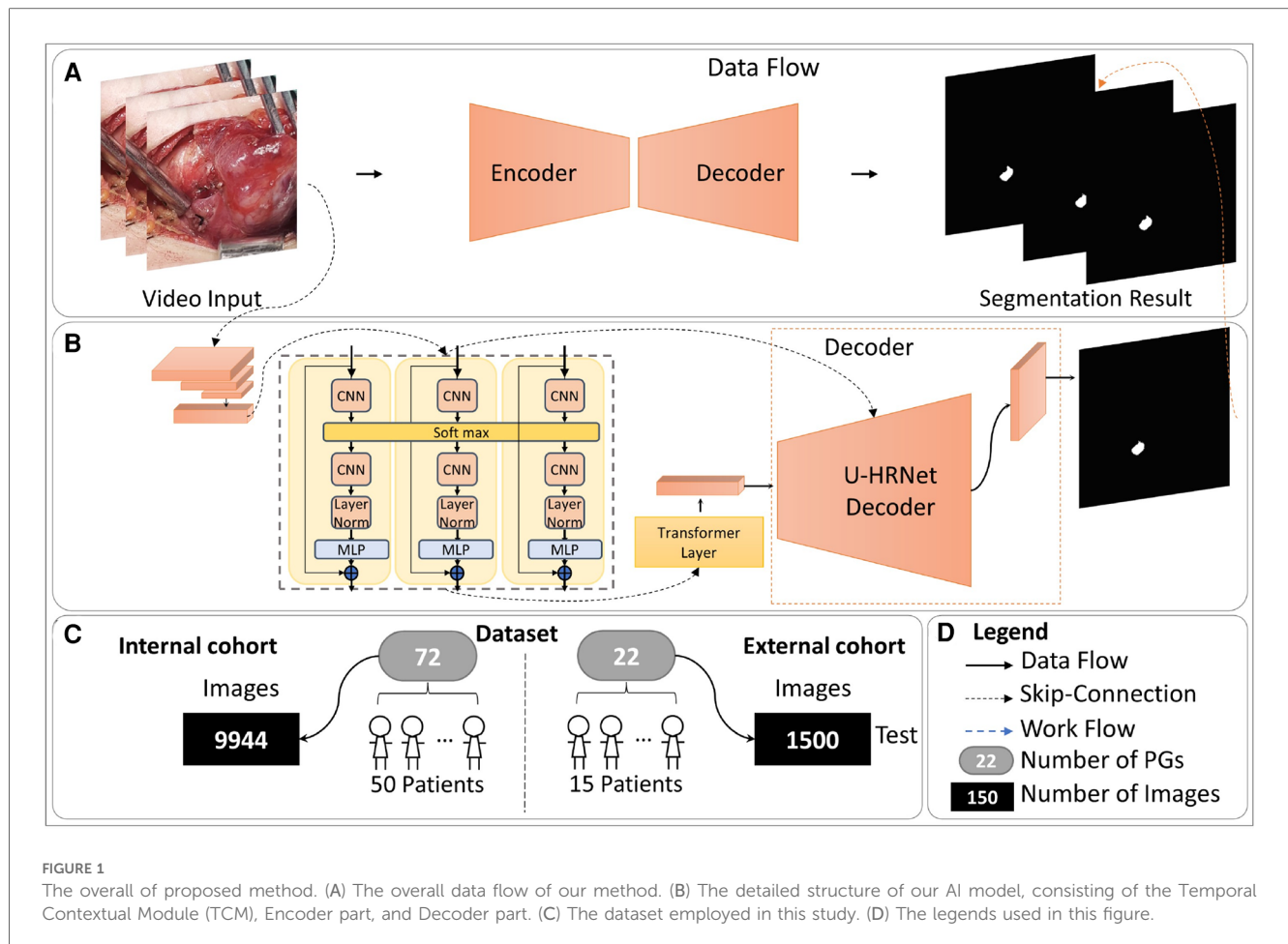
Skip-connections in U-shape net are crucial in multi-scale features transfer and fusion, we added a dynamic region-aware convolution (DRConv) module into skip-connection for preliminary extraction of multi-scale features (12). The specific module structure and our implementation details are shown in Figure 2.

To improve the communication of high-level semantic information and enhance the integration of low-level features, we modified the decoder structure based on Wang's work, as illustrated in Figure 3 (13). In contrast to the TransUnet decoder, the U-HRNet Decoder is more intricate, featuring additional feature transfer paths between adjacent layers and more fusion components. This redesign is geared towards achieving more accurate segmentation results. To align with reality, we implemented a post-processing method that excludes prediction results smaller than 50×50 pixels in size.

2.2 Patients and surgical technique

A prospective analysis was conducted on a total of 65 patients undergoing thyroid surgery at Shanghai Sixth People's Hospital from August 24, 2021, to June 17, 2023. Approval for this study was obtained from the Institutional Ethics Review Board and the Ethics Committee of Shanghai Sixth People's Hospital [Approval no. 2022-KY-178 (K)].

During surgery, a high-resolution camera was used to take videos of the wound surface, with the lens placed 15 cm away from the surgical field. Each patient underwent the recording of 1–2 videos. Each PG in videos was validated by the immune colloidal gold technique (ICGT) (5). A total of 116 videos, comprising 11,444 frames, were collected. Each video spanned 5–10 s, containing 50–100 frames. Within the internal cohort, 101 videos with 9,944 frames were meticulously labeled based on pathological results. In contrast, the external cohort consisted of



15 videos with 1,500 frames, serving solely for evaluation purposes and lacking specific segmentation labeling. Specific patient characteristics were shown in [Table 1](#).

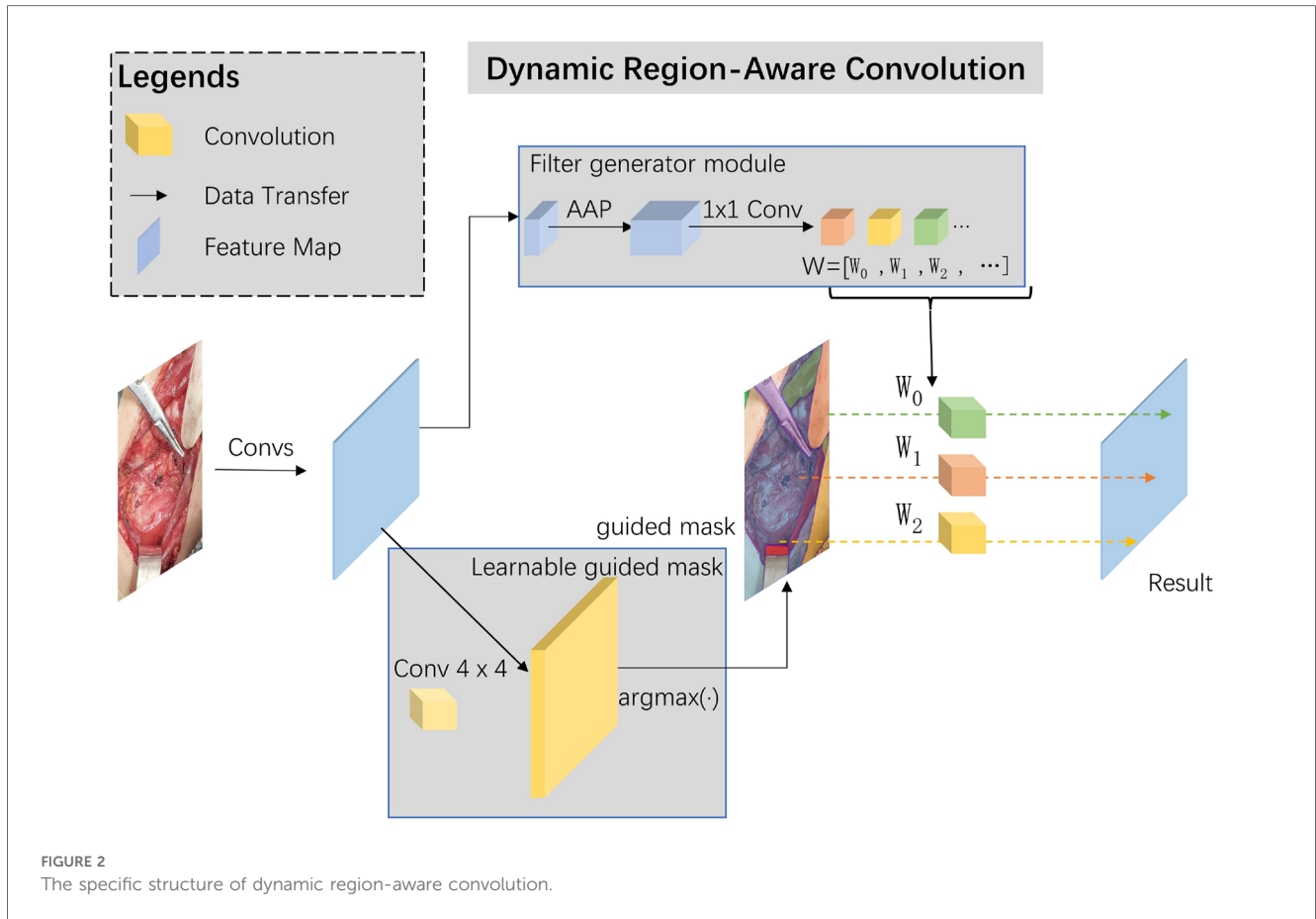
2.3 Dataset and comparison method

During this process, a senior surgeon with over 20 years of experience labeled all the PGs in videos using labelme software, including the location and contour of the PGs. The video data from 50 patients were allocated for internal training and validation sets, whereas the remaining data from 15 patients constituted the independent external validation cohort. The intersection over union (IoU) threshold selected for comparing prediction results with ground truth was set at 0.5. Performance evaluation of the deep learning model was based on the overlap in prediction masks generated by the model and those manually placed by the research team on each frame. Precision and recall were calculated for the study. Precision of deep learning models is similar to “positive predictive value,” and attempts to answer what proportion of positive classifications was actually correct. The specific calculations were placed in [Section 2.4](#). On the other hand, recall measures the model’s ability to detect positive samples, similar to “sensitivity,” and was calculated as the ratio between the number of positive samples correctly classified as positive to the total

number of positive samples. Two levels of surgeons evaluated the videos in the external validation cohort. Two levels of surgeons include 2 surgeons with more than 10 years of clinical experience in thyroidectomy and 2 surgeons with less than 10 years of clinical experience. They represented senior surgeons and junior surgeons respectively. All surgeons participating in the study had received specialist training in thyroid surgery. Our participating surgeons—both junior and senior—perform an average of 600 thyroid and parathyroid surgeries annually. This also includes the necessary training and assessments relevant to these procedures. Surgeons figured out the possible location of PGs in the video. Precision, recall, and accuracy were calculated. We compared the performance of the AI model and 2 levels of surgeons.

2.4 Evaluation metrics

For the evaluation of the AI model, we employed AI evaluation metrics. In this section, we introduced an evaluation metric specifically designed for PGs identification. Positive samples were calculated based on samples with an IoU threshold greater than 50% with the Ground Truth. Given that parathyroid glands are instances, segmentation results with an IoU of more than 50% were considered reasonable in the medical field. During the inference process, our method was evaluated in both the internal validation set and the



external test cohort. Precision, recall, and accuracy were abbreviated as PRE, REC, and ACC, respectively. True positive, false positive, false negative, and true negative were abbreviated as TP, FP, FN, and TN, respectively. Specific computations are shown below.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FN} + \text{FP} + \text{TN})$$

Dice coefficient and Jaccard coefficient is the most commonly used evaluation metric in medical segmentation field, and the formulas are as follows. While X represents the area of the Ground Truth, Y represents the area of Prediction results. Both Dice and Jaccard may evaluate how well the prediction results cover the Ground Truth. Jaccard is abbreviated as JAC.

$$\text{Dice} = 2 \times (|X \cap Y|) / (|X| + |Y|)$$

$$\text{Jaccard} = (|X \cap Y|) / (|X| + |Y| - |X \cap Y|)$$

For evaluation of video, we proposed a new coefficient for frame loss performance. Loss Frames represented the number of frames with no target. Total Frames represented the total number of frames. The

formula is as follows.

$$\text{FLR} = (\text{Loss Frames}) / (\text{Total Frames})$$

A χ^2 test was used to compare the PG recognition rate between the AI model and different groups of surgeons. Comparisons between groups were statistically processed by SPSS 26.0, and statistical significance was assigned for p values < 0.05 .

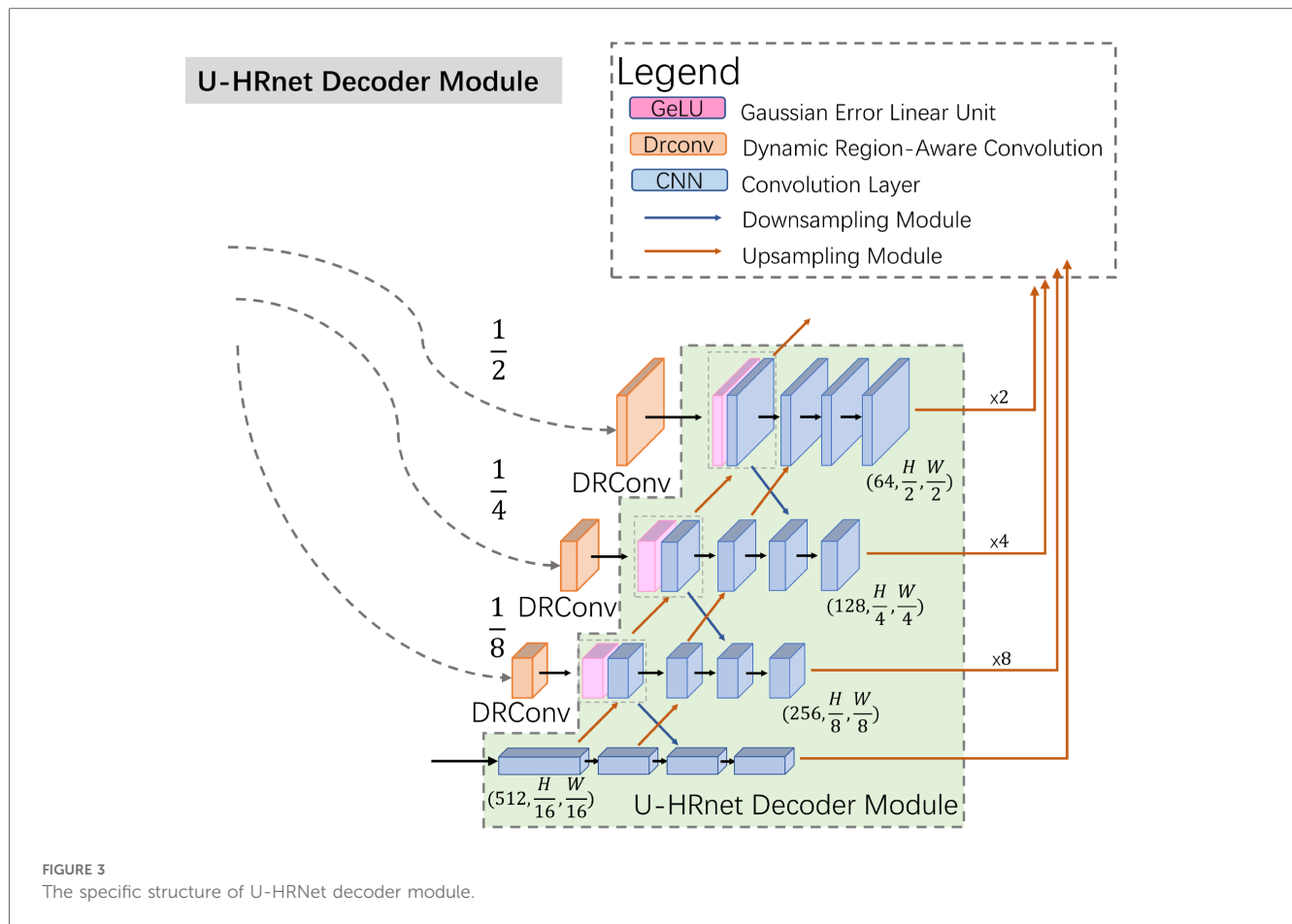
3 Results

3.1 Internal validation results

We conducted a comparative analysis of our method against three advanced deep-learning methods applied in the relevant medical field using our validation dataset. The results are presented in Table 2. As shown in Table 2, our proposed method outperformed the others in the validation dataset, indicating its superior performance. In internal validation cohort, we got 84.9% Precision and 81.3% Recall.

3.2 Comparison with surgeons

For external validation, our method was compared with two different levels of surgeons, junior surgeons and senior surgeons.



As the external cohort lacked specific labels, frames in the external validation dataset only had coarse labels, making accuracy evaluation metrics impractical for this cohort. Detailed

comparison results are presented in Table 3. In this comparison, our method achieved a precision of 89.5% and an accuracy of 70.8% in identifying PGs, outperforming all other surgeons. In comparison to all surgeons, our method achieved higher precision results as it tends to provide more accurate outcomes, while surgeons leaned towards offering more candidates to avoid missing PGs. Despite surgeons having more prior information, such as potential PG locations and their numbers, both the accuracy and precision of our method surpassed those of the surgeons. Comparison with senior surgeons revealed that surgeons with extensive experience identified more accurate PGs. The results suggest that the AI model can assist in providing more accurate PG results, thereby reducing the time and effort required for surgeons to identify PGs. Additionally, our method, based on a video recognition algorithm, demonstrates the ability to track PGs throughout the entire process of thyroidectomy, as depicted in Figure 4.

TABLE 1 Patient characteristics.

Parameter	Value
Age, year, median (IQR)	47 (21.5)
Sex, n (%)	
Female	47 (72.3)
Male	18 (27.7)
Diagnosis, n (%)	
Thyroid cancer	48 (73.8)
Thyroid nodule	12 (18.5)
Graves' disease	5 (7.7)
Procedure, n (%)	
Total thyroidectomy	28 (43.1)
Thyroid lobectomy	37 (42.1)
Hospital duration, day, median (IQR)	5 (3.5)
Operation time, min, median (IQR)	70 (35)
Parathyroid Glands	
Training	52
Internal Validation Cohort	20
Independent Test Cohort	22

Parathyroid glands represent the number of parathyroids used for model training.

In addition, we compared the differences between our model and two levels of surgeons in the external cohort. Our AI model demonstrated results comparable to those of senior surgeons, with no significant difference (senior surgeon 1: $\chi^2 = 0.989$, $p = 0.320$; senior surgeon 2: $\chi^2 = 1.373$, $p = 0.241$). And our model was superior to two junior surgeons, with a statistically significant difference (junior surgeon 1: $\chi^2 = 3.889$, $p = 0.048$; junior surgeon 2: $\chi^2 = 4.763$, $p = 0.029$).

TABLE 2 Result of proposed method.

Model	DICE	ACC	PRE	REC	hd_95	JAC	FLR
Our method	0.760	74.7%	80.2%	82.4%	28.44	0.769	11%
Video-TransUnet	0.710	70.1%	76.2%	71.4%	18.49	0.677	15%
Video-SwinUnet	0.754	73.6%	84.8%	80.4%	19.22	0.714	12%
TransUnet	0.705	69.4%	76.2%	74.3%	19.98	0.650	20%

TABLE 3 Result of proposed method.

Model	PRE	REC	ACC
Our method	89.5%	77.3%	70.8%
Junior surgeon 1	75.0%	81.8%	64.3%
Junior surgeon 2	80.0%	72.7%	64.0%
Senior surgeon 1	70.0%	63.6%	50.0%
Senior surgeon 2	84.6%	84.6%	45.8%

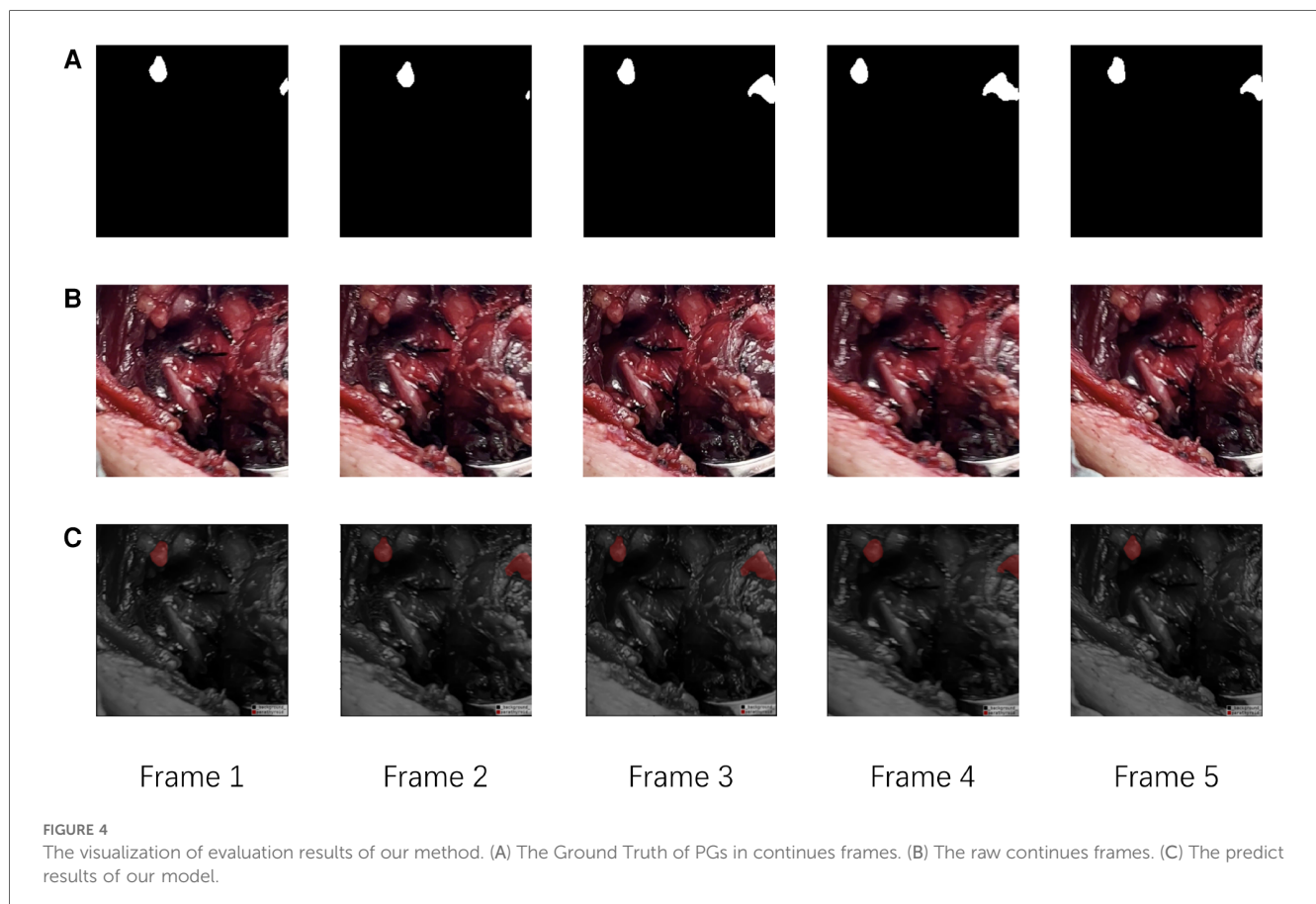
4 Discussion

In our retrospective study, we evidenced the proficiency of the proposed AI model in the recognition of PG in intraoperative videos. This model achieved a performance level demonstrating equivalence to that of two senior surgeons. A comparative analysis with multiple surgeons further solidified the potential of our presented method, indicating its application could bolster intraoperative recognition of PG. Ultimately, the implementation of this methodology could

precipitate a meaningful advancement in surgical accuracy, thereby substantially augmenting patient safety measures. To the best of our knowledge, this marks the inaugural instance of an AI model predicting the location and masks of PGs through the analysis of intraoperative videos in patients undergoing thyroidectomy.

AI in medicine holds the potential for expedited and standardized training. Kitaguchi et al. demonstrated the application of convolutional neural networks for automatic surgical skill assessment in laparoscopic colorectal surgery (14). Similarly, Wang and Fey described a deep learning framework for skill assessment in robotic surgery (15). As surgical procedures undergo increased digitalization, coupled with advancements in the automatic and real-time recognition of objects and tasks, the operating room stands as a focal point for progress. The newly developed tools have the capacity to enhance the performance and capabilities of surgical teams.

In the current medical landscape, video plays a crucial role in diagnostic solutions, spanning various applications such as



ultrasound, robotic surgery, and endoscopy (16). In contrast to static images, videos offer an effective means to comprehensively depict the entire process and intricacies of diagnosis or surgery. Doctors can grasp the situation and characteristics of lesions or surgical areas in multiple dimensions through time series data. Consequently, the study of video data in the medical field holds considerable significance. In 2022, Zeng et al. proposed Video-Transunet for assisted swallowing therapy in ultrasound video data. Based on this, Zeng's team proposed that Video-SwinUnet could also be applied to assist in swallowing therapy diagnosis in ultrasound videos (17). In the field of ultrasound diagnosis, Yeh et al. applied real-time object detection models to assess the recovery of the median nerve in ultrasound videos (18). Yu et al. also applied the U-Net network to the segmentation of surgical instruments in surgical video data for robotic surgery (19). The above research reflects that existing AI models can effectively assist in the automatic localization and segmentation of special targets in ultrasound videos or surgical videos.

Our proposed AI model demonstrated improved performance relative to several other advanced AI models. In evaluation of external cohort, our model equaled the performance of two experienced, senior surgeons with no significant difference ($p = 0.320$, $p = 0.241$; t -test). Notably, our model exhibited superior accuracy and precision, albeit with a reduced recall rate. Nevertheless, there are two primary limitations to our model. Firstly, due to the disparities in judgment between the AI model and the surgeons' recognition of PGs, our model might generate more false negatives than surgeons. This observation is evident from the comparison of recall rates between surgeons and our method. Secondly, as videos were recorded subsequent to the initial surgeon's PG identification, there exists the potential for false negatives or overlooked PGs. To mitigate these risks, an experienced surgeon, with over 20 years in the field, validated the PGs in the videos, with additional validation supplied by ICGT. These measures seek to minimize the likelihood of false negatives or missed PGs.

In practical application scenarios, due to the fixed number of PGs, incorrect judgment can easily lead to damage to the correct PG during surgery. Due to our method being more stable in determining the correct PG, it can provide surgeons with a certain degree of judgment advice. Moving forward, our ambition is to apply this method for real-time recognition of PG during surgical procedures. Concurrently, it shall function as an assistive tool for less experienced, junior surgeons in PG recognition, thereby minimizing the incidence of postoperative hypoparathyroidism. We anticipate that these efforts will play an instrumental role in enhancing patient prognosis.

In this retrospective study, we introduced a novel AI-based method to detect the shapes and localization of PGs in thyroid surgery, providing assistance to surgeons in PG detection. Comparative results with surgeons reveal that our method surpasses junior doctors in identifying PGs in intraoperative videos of thyroid surgery and performs at a level equivalent to that of senior surgeons. This provides valuable assistance for surgeons during thyroid surgery.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Institutional Ethics Review Board and the Ethics Committee of Shanghai Sixth People's Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

TS: Conceptualization, Formal Analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. FY: Data curation, Formal Analysis, Visualization, Writing – review & editing, Writing – original draft. JZ: Data curation, Supervision, Writing – review & editing. BW: Conceptualization, Funding acquisition, Project administration, Validation, Writing – review & editing. XD: Methodology, Project administration, Supervision, Writing – review & editing. CS: Project administration, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

This work is supported by the Interdisciplinary Program of Shanghai Jiao Tong University (No. YG2023LC10) and Cooperation Program of Ningbo Institute of Life and Health Industry, University of Chinese Academy of Sciences (No. 2019YJY0201).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher,

the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made

by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Xiang D, Xie L, Li Z, Wang P, Ye M, Zhu M. Endoscopic thyroidectomy along with bilateral central neck dissection (ETBC) increases the risk of transient hypoparathyroidism for patients with thyroid carcinoma. *Endocrine*. (2016) 53(3):747–53. doi: 10.1007/s12020-016-0884-y
- Choi JY, Lee KE, Chung KW, Kim SW, Choe JH, Koo do H, et al. Endoscopic thyroidectomy via bilateral axillo-breast approach (BABA): review of 512 cases in a single institute. *Surg Endosc*. (2012) 26(4):948–55. doi: 10.1007/s00464-011-1973-x
- Sitges-Serra A, Ruiz S, Girvent M, Manjon H, Duenas JP, Sancho JJ. Outcome of protracted hypoparathyroidism after total thyroidectomy. *Br J Surg*. (2010) 97(11):1687–95. doi: 10.1002/bjs.7219
- Khairy GA, Al-Saif A. Incidental parathyroidectomy during thyroid resection: incidence, risk factors, and outcome. *Ann Saudi Med*. (2011) 31(3):274–8. doi: 10.4103/0256-4947.81545
- Xia W, Zhang J, Shen W, Zhu Z, Yang Z, Li X. A rapid intraoperative parathyroid hormone assay based on the immune colloidal gold technique for parathyroid identification in thyroid surgery. *Front Endocrinol (Lausanne)*. (2020) 11:594745. doi: 10.3389/fendo.2020.594745
- Koimtzis G, Stefanopoulos L, Alexandrou V, Tteralli N, Brooker V, Alawad AA, et al. The role of carbon nanoparticles in lymph node dissection and parathyroid gland preservation during surgery for thyroid cancer: a systematic review and meta-analysis. *Cancers (Basel)*. (2022) 14(16):4016. doi: 10.3390/cancers14164016
- Solorzano CC, Thomas G, Baregamian N, Mahadevan-Jansen A. Detecting the near infrared autofluorescence of the human parathyroid: hype or opportunity? *Ann Surg*. (2020) 272(6):973–85. doi: 10.1097/SLA.0000000000003700
- Chen S, Lin Y, Li Z, Wang F, Cao Q. Automatic and accurate needle detection in 2D ultrasound during robot-assisted needle insertion process. *Int J Comput Assist Radiol Surg*. (2022) 17(2):295–303. doi: 10.1007/s11548-021-02519-6
- Zeng C, Yang X, Mirmehdi M, Gambaruto AM, Burghardt T, editors. *Video-TransUNet: temporally blended vision transformer for CT VFSS instance segmentation*. *International Conference on Machine Vision* (2022).
- Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, et al. TransUNet: transformers make strong encoders for medical image segmentation. *ArXiv [e-prints]*. arXiv:2102.04306 (2021).
- Yang X, Mirmehdi M, Burghardt T. Great ape detection in challenging Jungle Camera Trap Footage via attention-based spatial and temporal feature blending. *IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. (2019). p. 255–62. doi: 10.1109/ICCVW.2019.00034
- Chen J, Wang X, Guo Z, Zhang X, Sun J. Dynamic region-aware convolution. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. (2021). p. 8060–9. doi: 10.1109/CVPR46437.2021.00797
- Wang J, Long X, Chen G, Wu Z, Chen Z, Ding E. U-HRNet: delving into improving semantic representation of high resolution network for dense prediction. *arXiv [e-prints]*. arXiv:2210.07140 (2022).
- Kitaguchi D, Takeshita N, Matsuzaki H, Oda T, Watanabe M, Mori K, et al. Automated laparoscopic colorectal surgery workflow recognition using artificial intelligence: experimental research. *Int J Surg*. (2020) 79:88–94. doi: 10.1016/j.ijssu.2020.05.015
- Wang Z, Fey AM. SATR-DL: improving surgical skill assessment and task recognition in robot-assisted surgery with deep neural networks. *Annu Int Conf IEEE Eng Med Biol Soc*. (2018) 2018:1793–6. doi: 10.1109/EMBC.2018.8512575
- Kilincarslan O, Turk Y, Vargor A, Ozdemir M, Hassoy H, Makay O. Video gaming improves robotic surgery simulator success: a multi-clinic study on robotic skills. *J Robot Surg*. (2023) 17(4):1435–42. doi: 10.1007/s11701-023-01540-y
- Zeng C, Yang X, Smithard D, Mirmehdi M, Gambaruto AM, Burghardt T. Video-SwinUNet: Spatio-temporal Deep Learning Framework for VFSS Instance Segmentation. arXiv e-prints. (2023). arXiv:2302.11325.
- Yeh CL, Wu CH, Hsiao MY, Kuo PL. Real-time automated segmentation of median nerve in dynamic ultrasonography using deep learning. *Ultrasound Med Biol*. (2023) 49(5):1129–36. doi: 10.1016/j.ultrasmedbio.2022.12.014
- Yu L, Wang P, Yu X, Yan Y, Xia Y. A holistically-nested U-net: surgical instrument segmentation based on convolutional neural network. *J Digit Imaging*. (2020) 33(2):341–7. doi: 10.1007/s10278-019-00277-1