ORIGINAL RESEARCH

Laryngoscope
Investigative Otolaryngology

Anatomical recognition artificial intelligence for identifying the recurrent laryngeal nerve during endoscopic thyroid surgery: A single-center feasibility study

Yukio Nishiya MD, PhD [1,2] | Kazuto Matsuura MD, PhD [1] | Tateo Ogane [3] | Kazuyuki Hayashi MEng [3] | Yumi Kinebuchi MBA [3] | Hirotaka Tanaka MEng [4] | Wataru Okano MD, PhD [1] | Toshifumi Tomioka MD, PhD [1,3] | Takeshi Shinozaki MD, PhD [1] | Ryuichi Hayashi MD [1]

[1]Department of Head and Neck Surgery, National Cancer Center Hospital East, Chiba, Japan

[2]Department of Otolaryngology, The Jikei University School of Medicine, Tokyo, Japan

[3]Department of Medical Device Innovation, National Cancer Center Hospital East, Chiba, Japan

[4]Center for Promotion of Translational Research, National Cancer Center, Tokyo, Japan

Correspondence
Yukio Nishiya, Division of Head and Neck Surgery, National Cancer Center Hospital East, 6-5-1 Kashiwanoha, Kashiwa, Chiba, 277-8577, Japan.
Email: yukio@jikei.ac.jp

Funding information
Japan Society for the Promotion of Science

Abstract

Background: We investigate the feasibility of using artificial intelligence (AI) to identify the recurrent laryngeal nerve (RLN) during endoscopic thyroid surgery and evaluated its accuracy.

Methods: In this retrospective study, we develop an AI model using a dataset of endoscopic thyroid surgery videos, including hemithyroidectomy procedures performed between April 2019 and September 2023 at the National Cancer Center Hospital East, Chiba, Japan. Semantic segmentation deep learning methods were applied to analyze the endoscopic thyroid surgery videos.

Results: Forty endoscopic thyroid surgery videos, all in high definition or better quality, were analyzed. The Dice values were 0.351, 0.568, and 0.746 for the inferior thyroid artery, RLN, and trachea, respectively. Data augmentation was performed by cropping, standardizing, and resizing to reduce false positives and improve accuracy.

Conclusions: The AI model showed high recognition accuracy of the RLN and trachea. This method holds potential for assisting in future cervical gasless endoscopic surgeries.

KEYWORDS
artificial intelligence, deep learning, endoscopic thyroid surgery, recurrent laryngeal nerve, thyroidectomy

1 | INTRODUCTION

Conventional thyroid surgeries were performed through an open incision. However, in recent years, the use of remote-access endoscopic and robotic surgeries has increased.[1] The chest-breast approach,[2] bilateral axillo-breast approach,[3] retroauricular approach,[4] and transoral approach[5] are remote-access approaches that do not require a cervical incision. Thyroid disorders and tumors, which are prevalent among young women, have facilitated a high demand for aesthetic treatments. Remote-access approaches offer high cosmetic satisfaction as they avoid surgical scars in the neck area.[6] In contrast to open surgeries, good surgical videos can be easily recorded in endoscopic

surgeries. In contrast, endoscopic thyroid surgery takes longer than open surgery[7-9] and is reported to be associated with a higher frequency of recurrent laryngeal nerve (RLN) palsy,[10-12] particularly in facilities with limited experience. Furthermore, endoscopic thyroid surgeries have a long learning curve,[13-16] making them challenging to perform safely and accurately in less-experienced facilities.

The high frequency of complications and the extended recognition challenges may be attributed to anatomical features. Unlike thoracic and abdominal cavities, the neck region lacks a cavity structure. Thus, surgeons must first create a working space to perform an endoscopic procedure.[13] However, the space made by the surgeon is collapsible and unfamiliar, increasing the complexity of recognizing the dissection.

In areas related to the thyroid gland, many artificial intelligence (AI) diagnostic tools have been reported. Among these tools, the ultrasound is reported to be particularly useful in the diagnosis of benign and malignant.[17] Furthermore, it yields higher diagnostic accuracy than radiologists and ultrasonographers[18,19] and aids the assessment of genetic risk.[20] It has also been reported that ultrasound is useful for the diagnosis of benign and malignant disease.

However, only a few have directly catered to surgical applications. Although endoscopic thyroid surgery is a complex technique, it provides helpful surgical videos that are challenging to produce from open cervical incisional surgeries. This is particularly useful for developing new technologies and techniques. The rapid and accurate recognition of the RLN is an essential and challenging task because this nerve is the deepest and most critical anatomical feature in this field, particularly in gasless procedures. The trachea serves as a landmark for evaluating the estimated location of the RLN, and the inferior thyroid artery is a vessel that crosses the RLN; these are important structures for the surgical procedure. To determine whether AI can assist with endoscopic thyroid surgery safely without complications within a reasonable operating time at facilities with insufficient experience, we developed an AI model capable of anatomical recognition from endoscopic thyroid surgery videos. We further investigated the feasibility of this tool in assisting cervical endoscopic surgery.

2 | MATERIALS AND METHODS

2.1 | Study design and patient cohort

This retrospective study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines. The AI model was developed using a dataset containing all cases for which endoscopic thyroid surgery videos were available, including those of hemithyroidectomy performed between April 2019 and September 2023 at the National Cancer Center Hospital East (https://www.ncc.go.jp/en/ncce/index.html). Exclusion criteria included cases wherein (1) usable high-definition surgical videos or those of better quality were unavailable, (2) the RLN could not be preserved intraoperatively, and (3) the patient refused to participate.

In this study, the RLN is considered the most important anatomy. Cases where the RLN could not be preserved were excluded because the nerve would not be in its original correct position, which may reduce learning accuracy.

2.2 | Surgical approach

At our institution, we recommend endoscopic thyroid surgery when a tumor is less than 50 mm without invasion outside the thyroid capsule or metastasis to the lateral compartment. Further, we offer patients the option of endoscopic thyroid surgery for benign tumors smaller than 50 mm or thyroid cancers without invasion outside the thyroid capsule or metastasis to the lateral compartment. All endoscopic surgeries were performed using video-assisted neck surgery,[21] which uses an endoscopic approach through a small subclavian incision on the affected side. After separating the cervical subcutis, the cervical skin was lifted with a retractor, and a single camera port was placed in the cervical field. Then, the thyroid gland was approached from the lateral side by separating the sternohyoid and omohyoid muscles on the affected side, pulling the omohyoid and sternocleidomastoid muscles laterally with muscle hooks, and fixing the operative field. The sternothyroid muscle was ligated and resected from the thyroid gland. The perithyroidal vessels, including the superior and inferior thyroid arteries, were ligated using surgical clips or energy devices. Generally, identifying the RLN is crucial during surgery. Thus, intraoperative neuromonitoring (IONM) was intermittently used to positively identify the RLN and trace its path to the larynx to ensure its preservation. A standard OLYMPUS rigid endoscope with a diameter of 4 mm was used in all cases. Although this study includes surgical videos of three surgeons, the approach is the same for all.

2.3 | Annotation

The objects of recognition included the RLN, trachea, and inferior thyroid artery. The surgical videos were converted into MP4 format at a frame rate of 59.94 fps and annotated using the Computer Vision Annotation Tool (Ver 1.7.0, Intel Corporation). One of the endoscopic thyroid surgery surgeons performed all annotation tasks. The region of interest in each frame was precisely enclosed and individually labeled. Annotations were created from frames in which the target anatomical structures were clearly recognizable to new frames wherein they were only slightly recognizable retrospectively. Annotations were performed on a frame each time a forceps maneuver was performed around the recognition target, the camera position was changed, or the view of the surgical field was changed. The RLN and inferior thyroid arteries were annotated with the main trunk and branched only once each.

The inferior thyroid artery was annotated before ligation. This is because identifying the vessel before ligation is beneficial, while its relevance decreases after ligation, making subsequent annotation more challenging.
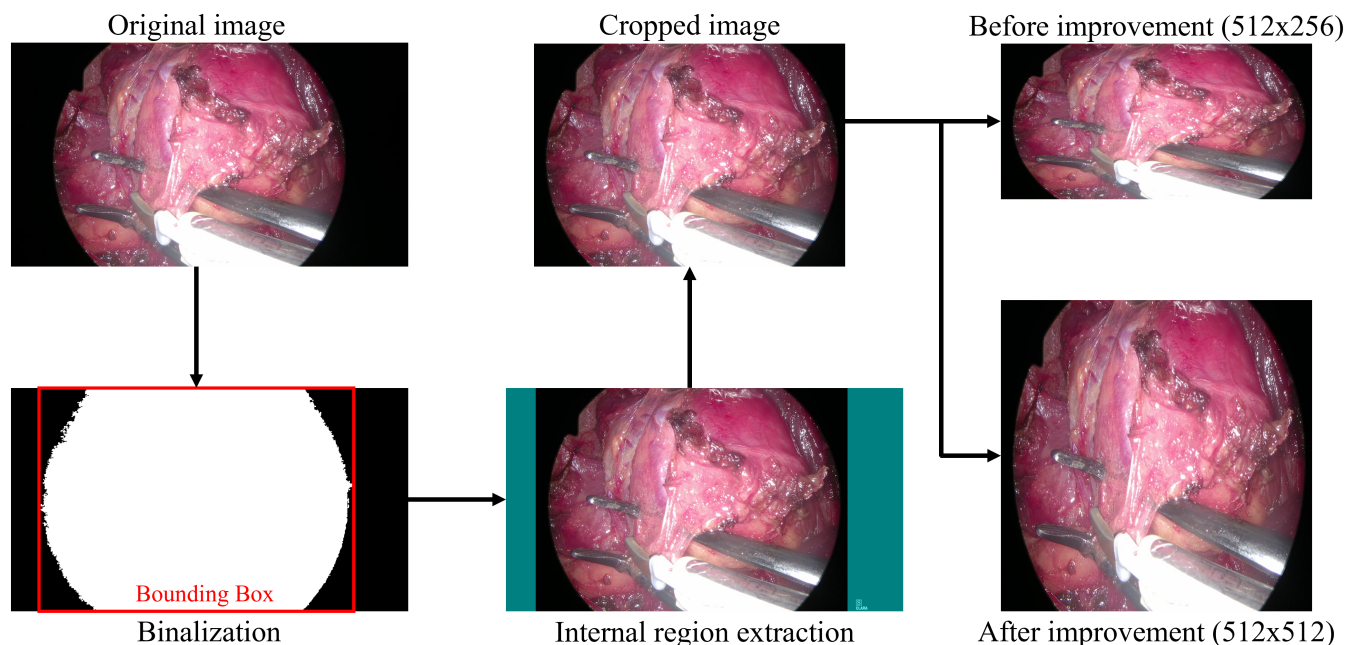
| Original image | Cropped image | Before improvement (512×256) |

Binalization | Internal region extraction | After improvement (512×512)

FIGURE 1 Cropped images with increased input size. Borders were identified, unnecessary areas were removed, sizes were enlarged, aspect was set to 1:1.
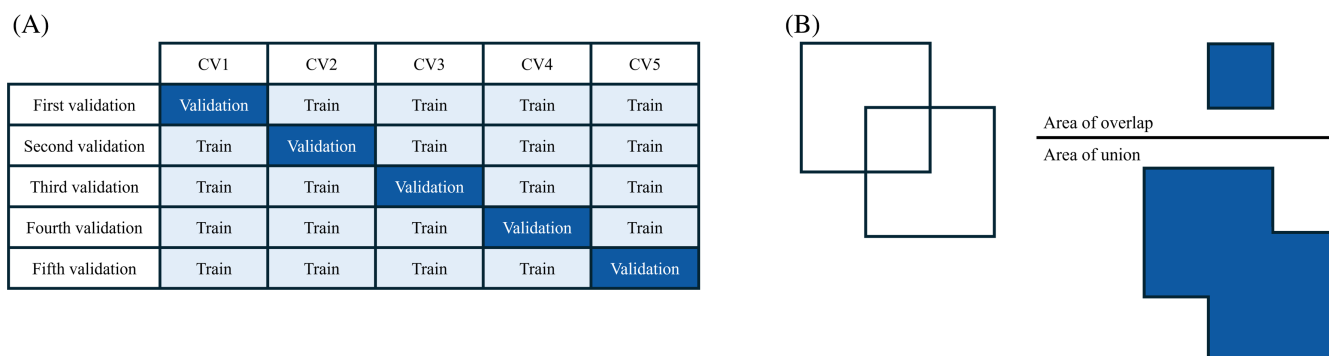

(A)

	CV1	CV2	CV3	CV4	CV5
First validation	Validation	Train	Train	Train	Train
Second validation	Train	Validation	Train	Train	Train
Third validation	Train	Train	Validation	Train	Train
Fourth validation	Train	Train	Train	Validation	Train
Fifth validation	Train	Train	Train	Train	Validation

(B)

Area of overlap / Area of union

FIGURE 2 Fivefold cross-validation and intersection-over-union. (A) fivefold cross-validation and (B) intersection-over-union.

2.4 | Ethics

The study protocol was approved by the institutional review board of the National Cancer Center Hospital East (registration no. 2022-358). Informed consent was obtained from patients in the form of opt-out decisions via the study website. Data from those who rejected participation were excluded. This study conformed to the provisions of the Declaration of Helsinki of 1964 (revised in Brazil in 2013).

2.5 | Model optimization and evaluation

Surgical videos were annotated for model learning and validation. Deep learning semantic segmentation tasks were performed using a convolutional neural network (CNN)-based approach. DeepLabv3+[22]

was used as the CNN model for the semantic segmentation task, and ImageNet (https://image-net.org/download.php) was used for pre-training. Additionally, the original and cropped images were learned and validated in each case. As shown in Figure 1, only the central region without an unnecessary black frame was extracted as a cropped image, whose input sizes were increased, and the aspect ratio was improved to 1:1. The images were standardized using mean subtraction = [133.579, 44.421, 64.584] and standardization = [47.897, 47.964, 45.970]. The trained model was evaluated in terms of precision, recall, and Dice score (F-measure) using the five-fold cross-validation method. As shown in Figure 2A, the data were divided into five groups, from CV1 to CV5. In particular, CV1 was used as the testing data in the first validation, and CV2–CV5 corresponded to the training data. This process was repeated five times, from CV1 to CV5. The averages of the results were used to determine the model accuracy.

TABLE 1 Training and validation dataset patient characteristic.

Cases	40
Age	61.6 (42–84)
Gender	
Male	16 (40%)
Female	24 (60%)
BMI	23.2 (17.6–31.6)
Surgical side	
Right	20 (50%)
Left	20 (50%)
Pathological diagnosis	
Benign	9 (23%)
Adenomatous nodule	5
Cyst	1
Follicular adenoma	3
Malignant	31 (76%)
Papillary carcinoma	27
Follicular carcinoma	3
Medullary carcinoma	1
Tumor size	20.1 (6–45)

Note: Average (range) or frequency (%).

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)}$$

where true positive (TP) refers to the annotated regions correctly predicted by the model, false positive (FP) refers to incorrect positive predictions, and false negative (FN) refers to correct regions that the model does not predict as a target region. The pixel-level results entail the computation of the TP, FP, and FN values for individual pixels within each image. The object-level results require calculating these metrics for each discernible object in an image. For the object-level results, TPs were defined with an intersection-over-union (IoU) equal to or greater than 0.1. As shown in Figure 2B, the IoU is an index that evaluates the amount of error between the object area predicted by the object detection model and the correct area. Specifically, the IoU is the overlapping area divided by the union area. Here, more than 10% overlap was present between the annotated correct region and the region predicted by the model. A confusion matrix[23] is a table frequently used to describe the performance of a model on a test dataset for which the true values are known. In our case, a confusion matrix was built using object-level results as per-pixel results possess considerably large values, thus increasing the complexity of visual interpretations.

TABLE 2 Details of 5-fold cross validation.

	CV1	CV2	CV3	CV4	CV5	Total
Cases	7	9	9	7	8	40
Frames	576	761	647	655	680	3319
Labels						
ITA	143	154	149	166	143	755
RLN	570	550	569	564	582	2835
Trachea	391	391	376	401	384	1943

Abbreviations: ITA, inferior thyroid artery; RLN, recurrent laryngeal nerve.

TABLE 3 Model accuracy based on average cross validation results.

	Pixel level		Object level	
	Original	Cropped	Original	Cropped
Dice				
ITA	0.249	0.351	0.161	0.232
Recurrent N	0.533	0.568	0.521	0.594
Trachea	0.715	0.746	0.423	0.479
Precision				
ITA	0.226	0.370	0.102	0.165
Recurrent N	0.504	0.556	0.393	0.484
Trachea	0.633	0.668	0.282	0.330
Recall				
ITA	0.278	0.335	0.384	0.393
Recurrent N	0.566	0.580	0.775	0.769
Trachea	0.821	0.845	0.840	0.873

Note: Original refers to the training results obtained using the original image of the surgery, and cropped refers to the training results obtained using the cropped image.
Abbreviations: ITA, inferior thyroid artery; RLN, recurrent laryngeal nerve.

3 | RESULTS

Eighty-two of the 134 patients who underwent hemithyroidectomy during the study period underwent open surgery, and 52 underwent endoscopic thyroid surgery. Of the 52 who underwent endoscopic thyroid surgery, 11 could not be considered owing to low image quality, and one case failed to preserve the RLN. Ultimately, as presented in Table 1, the dataset comprised 40 cases. None of the patients refused to participate. Half of the cases were right-sided, and more cases of malignant tumors than benign tumors were present. A total of 3319 frames and 5533 labels were annotated for the surgical videos of the 40 cases for training and validation. Table 2 demonstrates the details of fivefold cross-validation. Table 3 presents the validation results regarding the dice, precision, and recall scores. A higher accuracy was obtained with the cropped image at both pixel and object levels. Anatomically, the trachea exhibited the highest accuracy, followed by the RLN. Moreover, the inferior thyroid artery exhibited the lowest accuracy. Figure 3 shows the confusion matrix of
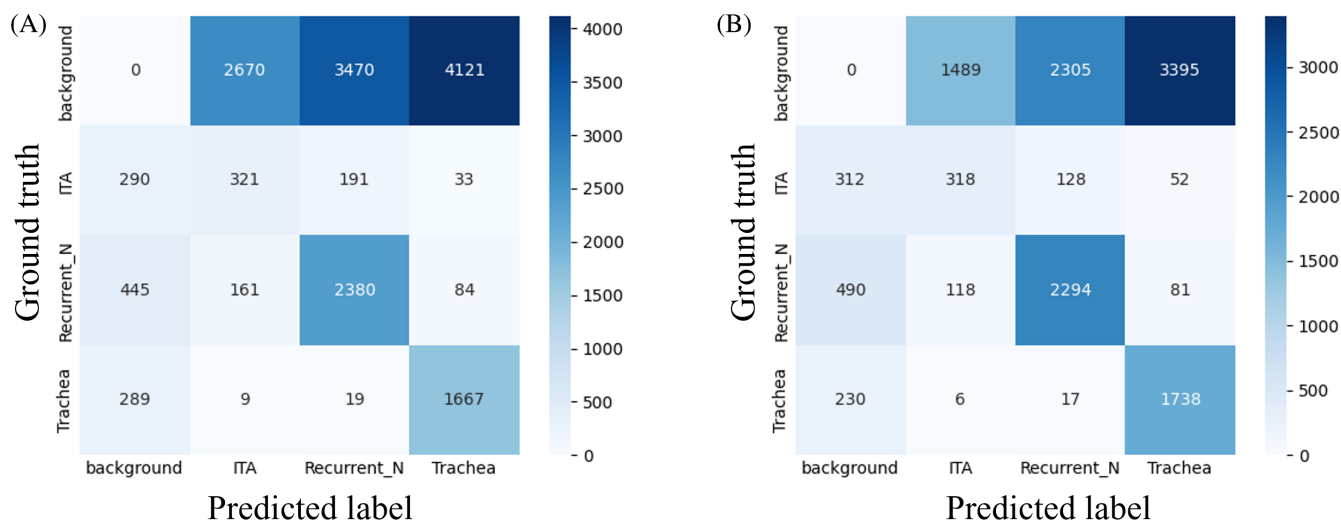
FIGURE 3 Confusion matrix. (A) Original and (B) cropped. A decrease in false positives was observed during cropped image learning.

FIGURE 4 Area distribution of (A) ground truth (GT) and (B) false positives (FPs). GT area peaked at 5–10k pixels; FPs were observed in smaller areas.

the training results obtained using the cropped and original images. Each label in the ground truth is an annotated region, and the background is an unannotated structure other than the target of this recognition. The part predicted by the AI model to be the target area for the background part is FPs, which decreased when cropped images were used. The ground truth and FP area distributions are illustrated in Figure 4. The peak area distribution of the ground truth was from 5000 to 10,000 pixels, and the FPs were mainly observed in small areas, especially those with less than 1000 pixels.

4 | DISCUSSION

We developed an AI model that recognizes anatomical images critical to surgical execution during endoscopic thyroid surgery. AI is increasingly being used for anatomical recognition in the abdominal and thoracic cavities, where minimally invasive surgeries are widely used.[24–26] However, only a few reports have focused on implementing AI models with gasless cervical endoscopic surgeries. Hence, this study is groundbreaking.

The three anatomies identified in this study included the RLN (i.e., the most important anatomical item to preserve during thyroid surgery), inferior thyroid artery, and trachea. Both the artery and trachea are crucial when searching for the RLN. The results show that the trachea exhibited the highest recognition accuracy, followed by the RLN, whereas the inferior thyroid artery had the lowest accuracy. The inferior thyroid artery is a vessel that must be ligated when the thyroid gland is removed, and it also helps identify the RLN. However, because the inferior thyroid artery is ligated intraoperatively, fewer training images are available than for the trachea and recurrent laryngeal nerves. ITA's recognition accuracy is low, so more annotation data is needed to improve it in the future. Additionally, because the

endoscopic cutting site is on the peripheral side of the inferior thyroid artery, many variations are present, and the area is small.[27] Therefore, achieving accurate results is challenging. Moreover, the trachea has no anatomical variation and serves as an orientation for surgery. Hence, AI models can yield highly accurate recognition. However, the accuracy of the RLN was intermediate in our results, and this may be because the RLN possesses a smaller area than the trachea. We posit that one of the reasons for the lower accuracy of the RLN compared to the trachea may be its deeper spatial location, which makes it more susceptible to even minor bleeding. Furthermore, this nerve is more variegated than the trachea but less variable and linear than the inferior thyroid artery. Dice results for the RLN and trachea differed between the per-pixel and per-object results possibly because the per-pixel results tended to have improved structures with larger areas. In our study, the trachea and RLN were recognized with good accuracy. The results are adequate compared to the accuracy reported in other surgical fields, such as the Dice pixel of the prostate (0.66)[24] azygos vein and vena cava (0.79), aorta (0.74),[25] and the RLN in the thoracic cavity (0.58).[26]

These findings suggest that AI-based image recognition technologies can be integrated in cervical endoscopic surgery using gasless methods. In the cervical region, Wang et al.[28] developed an AI model able to recognize parathyroid glands with high precision, recall rate, and F1 scores of 88.7%, 92.3%, and 90.5%, respectively. They also reported that the AI model outperformed junior and senior surgeons.[29] We suggest that, alongside the results of this study, the technology supporting thyroid endoscopic surgery will aid the implementation of this procedure safely and appropriately at any facility, thus increasing its utility. Furthermore, the application of this accurate anatomical recognition technology from the thyroid to the entire cervical region will contribute to endoscopic cervical dissections and eventually to the automation of surgery.

The primary surgical assistance device used in thyroid surgery is IONM.[30] IONM helps monitor and explore the RLN by inducing general anesthesia with an intubation tube using electrodes and stimulating the nerve with a probe or applying an automatic periodic stimulation electrode to the vagus nerve. Although the effectiveness of IONM for the early identification of the RLN has been emphasized, it does not significantly affect the rate of temporary or definitive RLN injury following thyroidectomy. However, its use is recommended in selected cases and for preventing bilateral palsy, as reported by Cozzi et al.[31]

By contrast, image recognition AI, which does not require preparation for monitoring and visually recognizes anatomies from endoscopic images or in combination with IONM, is expected to improve surgical support. In contrast, image recognition AI, which requires no preparation for monitoring and visually recognizes anatomy from endoscopic images, differs from IONM in terms of its operational mechanism. This may be problematic for future studies as it is expected to improve surgical support by AI alone or in combination with IONM.

Although this study utilized datasets from subclavian approach, different approaches (e.g., the bilateral axillo-breast approach endoscopic thyroidectomy (BABA-ET),[10] minimally invasive video-assisted thyroidectomy (MIVAT),[32] trans-oral endoscopic thyroidectomy through the vestibular approach (TOETVA),[5,33] and retroauricular approach endoscopic thyroidectomy (RAET)[4]) may involve different techniques for identifying the RLN and vary in difficulty. A systematic review by de Vries et al. reported that the median transient RLN injury rates were 3.7%, 2.5%, 4.0%, and 1.1% for BABA-ET, MIVAT, TOETVA, and RAET, respectively. However, these differences were not statistically significant. Therefore, the accuracy of RLN recognition across different techniques requires further study.

In this study, model accuracy was improved by increasing the training data and using cropped images to standardize and increase the size of the input images. These novel methods improved the results of previous attempts mainly by suppressing the FPs. This development is expected to be effective in cervical endoscopic surgeries that use a narrow and rigid speculum, as a black frame is present when a narrow endoscope is used.

Notably, several small-area FPs were observed because anatomical regions with minimal areas cannot be easily recognized in surgical videos and are rarely annotated. Therefore, the dissection of small areas might not be sufficiently learned or correctly predicted, or the AI's prediction might be accurate but not annotated, leading to a misclassification as an FP. The accuracy may be further improved by annotating small important areas and postprocessing to remove unimportant ones.

A limitation of this study is that our AI model was trained with single-center surgical videos, all corresponding to subclavian approaches, using images acquired with a 4-mm Olympus rigid speculum. Therefore, if a different scope is used, the accuracy of the results may vary owing to the different image quality obtained. Additionally, all annotations were performed by a single surgeon, addressing variability between annotators. However, there is a possibility of subjective bias. If multiple trained annotators could have performed the annotation, it may have provided a more robust ground truth. We lack a sufficient dataset to separately examine the difference in recognition accuracy for the left and right sides. However, the positional relationship between the inferior thyroid artery and RLN differs slightly between the left and right sides, thus requiring further investigation into the difference in recognition accuracy. Annotation of 40 cases may be a small number of cases for AI learning. In the future, whether the developed AI model can recognize other surgical environments and approaches with high accuracy should be determined.

5 | CONCLUSION

In this study, an image recognition AI model was developed for endoscopic thyroid surgery using a gasless method. AI image recognition technology may now be applied to cervical endoscopic surgeries because the recurrent laryngeal nerves and trachea were accurately identified using this method.

Yukio Nishiya has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

ORCID

Yukio Nishiya https://orcid.org/0000-0003-1666-3442
Wataru Okano https://orcid.org/0000-0003-3915-4229
Toshifumi Tomioka https://orcid.org/0000-0003-0406-3976
Takeshi Shinozaki https://orcid.org/0000-0001-6762-8490

REFERENCES

1. Lu Q, Zhu X, Wang P, Xue S, Chen G. Comparisons of different approaches and incisions of thyroid surgery and selection strategy. *Front Endocrinol (Lausanne)*. 2023;14:1166820. doi:10.3389/fendo.2023.1166820
2. Zhang Z, Xia F, Li X. Ambulatory endoscopic thyroidectomy via a chest-breast approach has an acceptable safety profile for thyroid nodule. *Front Endocrinol (Lausanne)*. 2021;12:795627. doi:10.3389/fendo.2021.795627
3. Johri G, Chand G, Mishra A, et al. Endoscopic versus conventional thyroid surgery: a comparison of quality of life, cosmetic outcomes, and overall patient satisfaction with treatment. *World J Surg*. 2020;44:4118-4126. doi:10.1007/s00268-020-05732-7
4. Lee DY, Baek S-K, Jung K-Y. Endoscopic thyroidectomy: retroauricular approach. *Gland Surg*. 2016;5:327-335. doi:10.21037/gs.2015.10.01
5. Anuwong A. Transoral endoscopic thyroidectomy vestibular approach: a series of the first 60 human cases. *World J Surg*. 2016;40(3):491-497. doi:10.1007/s00268-015-3320-1
6. Li X, Ding W, Zhang H. Surgical outcomes of endoscopic thyroidectomy approaches for thyroid cancer: a systematic review and network meta-analysis. *Front Endocrinol*. 2023;14:1256209. doi:10.3389/fendo.2023.1256209
7. Xu T, Qin X, Zhang Y, et al. A prospective study comparing the gasless endoscopic thyroidectomy trans-axillary approach to conventional open thyroidectomy: health and quality of life outcomes. *Surg Endosc*. 2024;38:1995-2009. doi:10.1007/s00464-024-10689-y
8. de Vries LH, Aykan D, Lodewijk L, Damen JAA, Borel Rinkes IHM, Vriens MR. Outcomes of minimally invasive thyroid surgery–a systematic review and meta-analysis. *Front Endocrinol*. 2021;12:719397. doi:10.3389/fendo.2021.719397
9. Ge J-N, Yu ST, Tan J, et al. A propensity score matching analysis of gasless endoscopic transaxillary thyroidectomy with five-settlement technique versus conventional open in patients with papillary thyroid microcarcinoma. *Surg Endosc*. 2023;37:9255-9262. doi:10.1007/s00464-023-10473-4
10. Choi JY, Lee KE, Chung K-W, et al. Endoscopic thyroidectomy via bilateral axillo-breast approach (BABA): review of 512 cases in a single institute. *Surg Endosc*. 2012;26:948-955. doi:10.1007/s00464-011-1973-x
11. Chung YS, Choe J-H, Kang KH, et al. Endoscopic thyroidectomy for thyroid malignancies: comparison with conventional open thyroidectomy. *World J Surg*. 2007;31:2302-2306; Discussion 2307–2308. doi:10.1007/s00268-007-9117-0
12. Mo L-L, Meng F-L, Yang Z-Q, Hou L-M, Fang F. Comparison of safety and efficacy between total endoscopic resection and conventional open surgery for malignant thyroid tumors: a meta-analysis. *Transl Cancer Res*. 2020;9:2865-2874. doi:10.21037/tcr.2020.02.29
13. Imran M, Mehmood Z, Baloch MN, Altaf S. Endoscopic thyroid lobectomy vs conventional open thyroid lobectomy. *Pak J Med Sci*. 2020;36:831-835. doi:10.12669/pjms.36.4.1604
14. Shimizu K, Shimizu K, Okamura R, et al. Video-assisted neck surgery (VANS) using a gasless lifting procedure for thyroid and parathyroid diseases: "The VANS method from A to Z". *Surg Today*. 2020;50:1126-1137. doi:10.1007/s00595-019-01908-4
15. Nagaoka R, Sugitani I, Kazusaka H, et al. Learning curve for endoscopic thyroidectomy using video-assisted neck surgery: retrospective analysis of a surgeon's experience with 100 patients. *J Nippon Med Sch*. 2022;89:277-286. doi:10.1272/jnms.JNMS.2022_89-302
16. Capponi MG, Bellotti C, Lotti M, Ansaloni L. Minimally invasive video-assisted thyroidectomy: ascending the learning curve. *J Minim Access Surg*. 2015;11:119-122. doi:10.4103/0972-9941.153808
17. Zhu P-S, Zhang Y-R, Ren JY, et al. Ultrasound-based deep learning using the VGGNet model for the differentiation of benign and malignant thyroid nodules: a meta-analysis. *Front Oncol*. 2022;12:944859. doi:10.3389/fonc.2022.944859
18. Chen Y, Gao Z, He Y, et al. An artificial intelligence model based on ACR TI-RADS characteristics for US diagnosis of thyroid nodules. *Radiology*. 2022;303:613-619. doi:10.1148/radiol.211455
19. Wang J, Jiang J, Zhang D, et al. An integrated AI model to improve diagnostic accuracy of ultrasound and output known risk features in suspicious thyroid nodules. *Eur Radiol*. 2022;32:2120-2129. doi:10.1007/s00330-021-08298-7
20. Daniels K, Gummadi S, Zhu Z, et al. Machine learning by ultrasonography for genetic risk stratification of thyroid nodules. *JAMA Otolaryngol Head Neck Surg*. 2020;146:36-41. doi:10.1001/jamaoto.2019.3073
21. Shimizu K, Akira S, Tanaka S. Video-assisted neck surgery: endoscopic resection of benign thyroid tumor aiming at scarless surgery on the neck. *J Surg Oncol*. 1998;69:178-180. doi:10.1002/(SICI)1096-9098(199811)69:33.0.CO;2-9
22. Chen L-C, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder–decoder with atrous separable convolution for semantic image segmentation. ECCV 2018: Computer Vision, pp. 833–851. 2018.
23. Powers DMW. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness, and correlation. *J Mach Learn Technol*. 2011;2:37-63.
24. Kitaguchi D, Takeshita N, Matsuzaki H, et al. Computer-assisted real-time automatic prostate segmentation during TaTME: a single-center feasibility study. *Surg Endosc*. 2021;35:2493-2499. doi:10.1007/s00464-020-07659-5
25. den Boer RB, Jaspers TJM, de Jongh C, et al. Deep learning-based recognition of key anatomical structures during robot-assisted minimally invasive esophagectomy. *Surg Endosc*. 2023;37:5164-5175. doi:10.1007/s00464-023-09990-z
26. Sato K, Fujita T, Matsuzaki H, et al. Real-time detection of the recurrent laryngeal nerve in thoracoscopic esophagectomy using artificial intelligence. *Surg Endosc*. 2022;36:5531-5539. doi:10.1007/s00464-022-09268-w
27. Bliss RD, Gauger PD, Delbridge LW. Surgeon's approach to the thyroid gland: surgical anatomy and the importance of technique. *World J Surg*. 2000;24:891-897. doi:10.1007/s002680010173
28. Wang B, Zheng J, Yu JF, et al. Development of artificial intelligence for parathyroid recognition during endoscopic thyroid surgery. *Laryngoscope*. 2022;132:2516-2523. doi:10.1002/lary.30173
29. Wang B, Yu JF, Lin SY, et al. Intraoperative AI-assisted early prediction of parathyroid and ischemia alert in endoscopic thyroid surgery. *Head Neck*. 2024;46:1975-1987. doi:10.1002/hed.27629
30. Dralle H, Sekulla C, Lorenz K, Brauckhoff M, Machens A. Intraoperative monitoring of the recurrent laryngeal nerve in thyroid surgery. *World J Surg*. 2008;32:1358-1366. doi:10.1007/s00268-008-9483-2

31. Cozzi AT, Ottavi A, Lozza P, et al. Intraoperative neuromonitoring does not reduce the risk of temporary and definitive recurrent laryngeal nerve damage during thyroid surgery: a systematic review and meta-analysis of endoscopic findings from 73,325 nerves at risk. *J Pers Med.* 2023;13:1429. doi:10.3390/jpm13101429

32. Pisanu A, Podda M, Reccia I, Porceddu G, Uccheddu A. A systematic review with meta-analysis of prospective randomized trials comparing minimally invasive video-assisted thyroidectomy (MIVAT) and conventional thyroidectomy (CT). *Langenbecks Arch Surg.* 2013;398: 1057-1068. doi:10.1007/s00423-013-1125-y

33. Akritidou E, Douridas G, Spartalis E, Tsourouflis G, Dimitroulis D, Nikiteas NI. Complications of trans-oral endoscopic thyroidectomy vestibular approach: a systematic review. *In Vivo.* 2022;36(1):1-12. doi:10.21873/invivo.12671