




An artificial intelligence-based recognition model of colorectal liver metastases in intraoperative ultrasonography with improved accuracy through algorithm integration

Maho Takayama | Kyoji Ito | Kenji Karako | Yuichiro Mihara | Shu Sasaki | Akihiko Ichida | Takeshi Takamoto | Nobuhisa Akamatsu  | Yoshikuni Kawaguchi  | Kiyoshi Hasegawa 

Hepato-Biliary-Pancreatic Surgery Division, Artificial Organ and Transplantation Division, Department of Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Correspondence

Kiyoshi Hasegawa, Hepato-Biliary-Pancreatic Surgery Division, Artificial Organ and Transplantation Division, Department of Surgery, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.
Email: kihase-tyk@umin.ac.jp

Funding information

Japan Society for the Promotion of Science, Grant/Award Number: 20K20214; Japan Surgical Society, Grant/Award Number: the 106th annual congress of JSS Memorial Surgical

Abstract

Background/Purpose: Contrast-enhanced intraoperative ultrasonography (CE-IOUS) is crucial for detecting colorectal liver metastases (CLM) during surgery. Although artificial intelligence shows potential in diagnostic systems, its application in CE-IOUS is limited.

Methods: This study aimed to develop an automatic tumor detection model using Mask region-based convolutional neural network (Mask R-CNN) for CE-IOUS images. CE-IOUS videos of the CLM from 121 patients were collected, generating ground truth data. A total of 2659 images were obtained. Two models were developed: the basic recognition model (BRM), which was trained on CE-mode images, and the subtraction model (SM), which used images created by a subtraction algorithm that highlighted the differences in pixel values between the basic-mode and CE-mode images. The subtraction algorithm focuses on echogenicity differences. These two models were combined into a combination model (CM), which assessed outcomes using the prediction probabilities from both models.

Results: The optimal epochs were determined by the maximum area under the curve (AUC), and the thresholds were calculated accordingly. BRM, SM, and CM achieved 89.4%, 86.6%, and 96.5% accuracy, respectively. CM outperformed the individual models, achieving an AUC of 0.99.

Conclusions: A novel automated recognition model was developed for accurate CLM detection in CE-IOUS by integrating image- and algorithm-based models.

KEYWORDS

colorectal cancer, computer-assisted diagnosis, liver resection, machine learning, ultrasonography

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Journal of Hepato-Biliary-Pancreatic Sciences* published by John Wiley & Sons Australia, Ltd on behalf of Japanese Society of Hepato-Biliary-Pancreatic Surgery.

1 | INTRODUCTION

Recent technical improvements in diagnostic modalities and surgical techniques have allowed aggressive resection of metastatic liver tumors. Colorectal liver metastases (CLM) are the most frequent metastatic liver tumors, and complete resection can improve the prognosis.

Intraoperative ultrasonography (IOUS) is an essential tool for detecting liver tumors during hepatectomy. In particular, contrast-enhanced IUOS (CE-IOUS) using perflubutane (Sonazoid; GE Healthcare, Oslo, Norway), which is a recent ultrasound contrast agent useful for minimally invasive tumor detection, is highly effective for diagnosing and identifying CLM during surgery.¹ However, accurate identification of lesions is sometimes difficult due to technical difficulties² and difficulty in distinguishing between liver tumors and vascular vessels. Moreover, due to its clinicopathological characteristics, CLM often manifests as multiple lesions within the liver, making it challenging to detect all lesions thoroughly and within the limited time available in clinical practice.

In recent years, the use of artificial intelligence (AI) technology has led to the development of diagnostic support systems in various medical fields.^{3,4} Several AI models for liver tumor diagnosis have been reported, and lesions can be accurately detected by combining AI models.^{5–7} Previous studies have attempted to build better AI systems by adjusting various learning methods.

However, few studies have combined machine learning with algorithmic processing based on the characteristics of tumor images to improve the diagnostic performance. In addition, few reports exist on the application of AI systems in IOUS, especially CE-IOUS.

The purpose of this study was to apply AI to CE-IOUS images and develop an automatic recognition system for CLM. We not only used CE-IOUS images to learn tumor images but also combined algorithmic processing based on the characteristics of tumor images and machine learning. Subsequently, a predictive model that combined the results of the two different models was created using machine learning, and we compared the accuracy of both methods.

2 | METHODS

2.1 | Study design

Our purpose was to accurately identify the location of tumors in ultrasound images in real time. To achieve this, we created a model for tumor detection based on the mask region-based convolutional neural network (Mask

R-CNN).⁸ Mask R-CNN is an object detection method ((2018) FAIR's research platform for object detection research, implementing popular algorithms such as Mask R-CNN and RetinaNet.⁹: facebookresearch/Detection. Facebook Research) and a deep neural network was used for object segmentation. Mask R-CNN is a modified model based on the Faster R-CNN, which predicts the object's class, refines the bounding box, and creates a pixel-level mask for the object.¹⁰ Using Mask R-CNN, we created two different models; the two models were then integrated to improve the reliability of the predictions and reduce false detections.

2.2 | Data collection and preparation

2.2.1 | Collection of CE-IOUS videos for CLM

Liver resection for CLM performed at the University of Tokyo Hospital between January 2020 and August 2022 were included. Surgical approaches included open, laparoscopic, and robotic-assisted resection. Cases of multiple tumors as well as single tumors were included in the analysis. This study was approved by the appropriate institutional review board of the Graduate School of Medicine and Faculty of Medicine, The University of Tokyo (no. 2019166NI) and informed consent was obtained in the form of opt-out on the website. In all cases, CE-IOUS was performed before hepatic resection to detect tumor location. To obtain contrast enhancement images, perflubutane (Sonazoid) was administered (0.5 mL intravenously). The mechanism of action of perflubutane in CE-IOUS is described in Supplementary Method 1—Data S1.

CLM detection was performed by comparing the B- and CE-mode images. Tumor screening was conducted 10–15 min after perflubutane administration (Kupffer phase). The CE-IOUS images were preserved as videos.

2.2.2 | Creation of ground truth data

In the collected CE-IOUS videos, tumors with characteristic ultrasound findings and pathologically confirmed CLM were used for training and testing. The ultrasound findings for CLM are shown in Figure 1a. A CLM is typically visualized as a hyperechoic nodule on a B-mode image and as a contrast defect on a CE-mode image.

First, the scenes in the CE-mode videos where the tumor appeared hypoechoic were selected and cropped as still images (Figure 1b, left). In order to develop a system

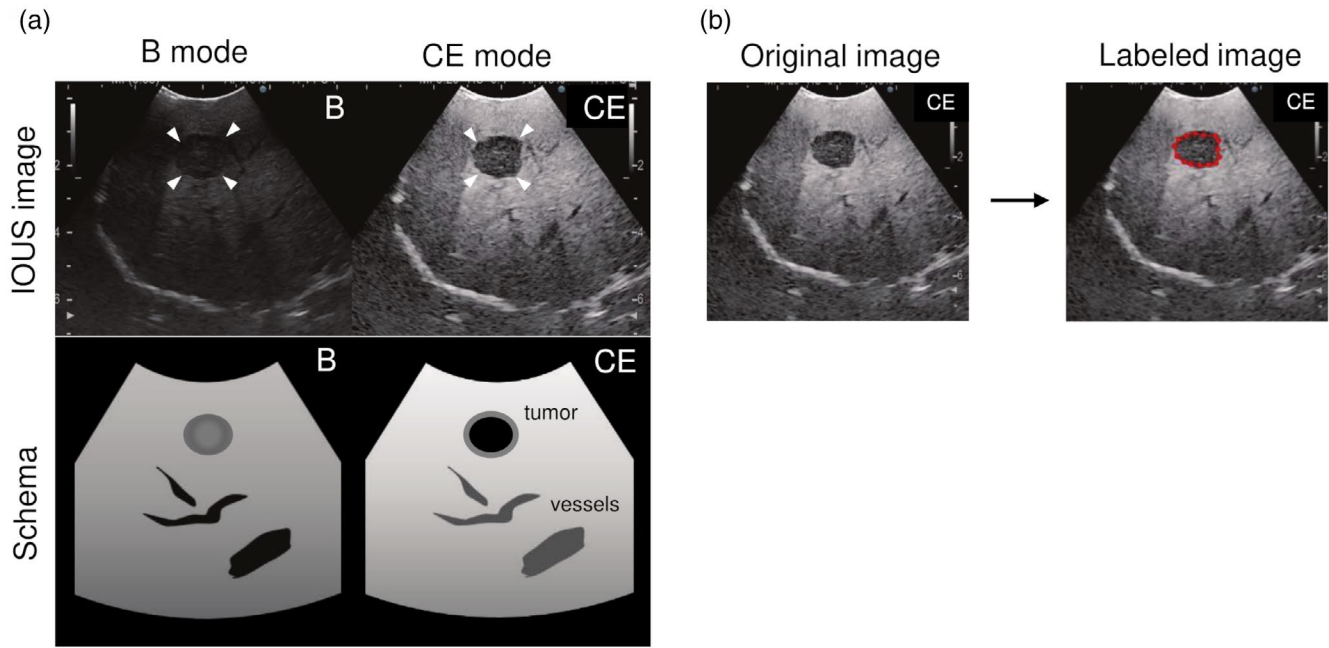


FIGURE 1 Ultrasonographic findings of typical CLM and labeling of the tumor. (a) IOUS image (upper row) and schema (lower row) of the CLM (arrowhead) on B-mode- and CE-mode images. The tumor location is indicated by arrowheads. A CLM is typically visualized as a hyperechoic nodule in B mode and as a hypoechoic nodule in CE mode. (b) The tumors in the CE-mode image are labeled by enclosing the tumor with dots and lines. B mode, basic mode; CE mode, contrast-enhanced mode; CLM, colorectal liver metastasis; IOUS, intraoperative ultrasonography.

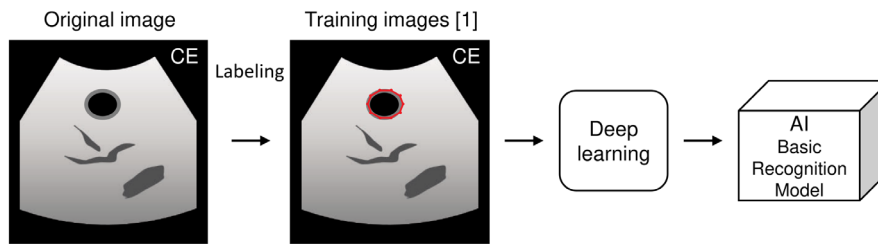


FIGURE 2 Construction of BRM. The flow of constructing the BRM is shown in the schema. CLMs on the CE-mode images are labeled as the training data, and machine learning is then conducted using the labeled training images. AI, artificial intelligence; BRM, basic recognition model; CE mode, contrast-enhanced mode; CLM, colorectal liver metastases.

that could accurately respond to the differences in tumor localization, these images included various segments of the liver, deep and surface localization, proximity, and remoteness to blood vessels. These images that were missing or did not depict the entire tumor were excluded. Tumors in the still images in CE mode were labeled by surrounding them with dots and lines (Figure 1b, right). When multiple tumors were depicted on a single ultrasound screen, each tumor was labeled independently. The tumors were labeled by two experienced hepatobiliary and pancreatic surgeons (M.T and Y.M). These labeled images were divided into typical images (tumor visible as a hypoechoic area) and atypical images (small tumors (<1 cm), tumors with unclear hypoechoic areas). These images were used in the training and test data as images of the tumors. Next, scenes from the CE-mode video that did not show the

tumor were selected and cropped into still images. These were used in the test data without tumor images.

2.2.3 | Creation of basic recognition model (BRM)

We constructed an object detection model for tumor identification using the CE-mode training images (right half of the ultrasound images) as the input data (Figure 2). The model employed a Mask R-CNN.⁸ For actual implementation, we used PyTorch 2.1^{11,12} and performed transfer learning aimed at tumor detection using MaskRCNN_ResNet50_FPN_Weights.COCO_V1 was pretrained on the COCO dataset,^{13,14} available through the torchvision library.¹⁵ The training was performed with

settings that included a batch size of four, learning rate of 0.001, 50 epochs, and the Adam optimizer.¹⁶ The training data were enhanced by applying random noise for data augmentation. Through this training, we developed a model capable of detecting tumors and estimating the mask regions of objects from images in the CE mode.

2.2.4 | Creation of subtraction model (SM)

It is generally difficult to distinguish a longitudinally sliced Glissonian sheath from a small tumor on ultrasound images. In the B mode, the longitudinally sliced

Glissonian sheath has a hypoechoic lumen and a hyperechoic perivascular area. The CLM has a hyperechoic interior and hypoechoic exterior. On the other hand, in the CE mode, the lumen of the Glissonian vessels is isoechoic or slightly hyperechoic compared to the B mode due to microbubbles. This is in contrast to the inside of the tumor, which is hypoechoic in CE mode than in B mode (Figure 3a). Using this difference, an algorithm to distinguish between vessels and tumors was incorporated into the images before deep learning to increase the recognition rate of Glissonian sheaths and small tumors.

The B-mode and labeled CE-mode images were used to create the BRM. We created new images (subtraction

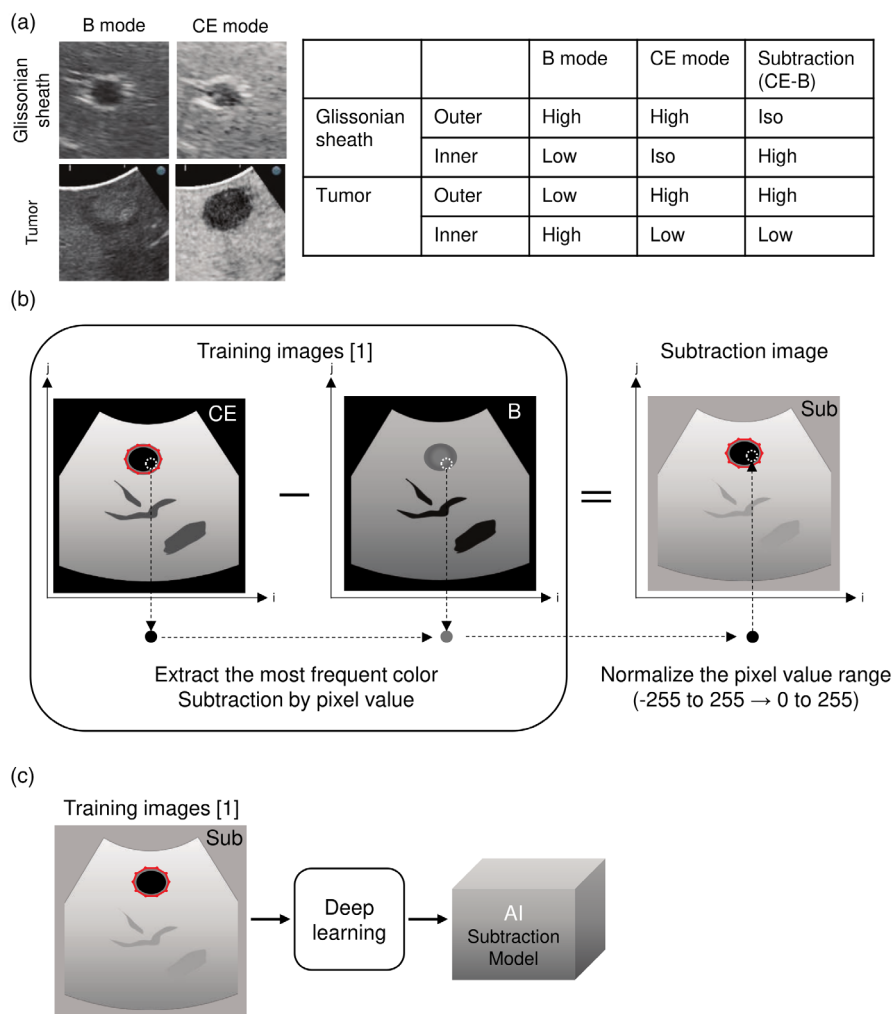


FIGURE 3 Construction of SM. (a) A comparison of the B- and CE-mode ultrasound images between the Glissonian sheath image (upper row) and CLM (lower row). The ultrasonographic intensities of the Glissonian sheath and CLM stratified by the inner and outer layers in B mode, CE mode, and subtraction images (CE-B) are shown in the table. (b) Creation of subtraction images. The B-mode and labeled CE-mode images were used to create subtraction images by subtracting the pixel values of the B-mode images from those of the CE-mode images. The most frequent pixel values from small circular areas for each coordinate i and j in the B- and CE-mode images were subtracted, and the resulting values were normalized for imaging purposes. (c) Construction of SM. The flow of constructing the SM is shown in the schema. CLMs on the subtracted images are labeled as the training data, and machine learning is then conducted using the labeled training images. B mode, basic mode; CE mode, contrast-enhanced mode; CLM, colorectal liver metastases; High, hyperechoic; Iso, isoechoic; Low, hypoechoic; SM, subtraction model; Sub, subtraction image.

images) by subtracting the pixel values of the B-mode images from those of the CE-mode images (Figure 3b). Specifically, each coordinate i and j in the B- and CE-mode images were converted to grayscale, and a histogram was calculated for a small circular area with a radius of 10 pixels. The most frequent pixel values (Mode) were extracted from the two histograms for both $\text{Mode}_{\text{B-mode}}$ and $\text{Mode}_{\text{CE-mode}}$. The subtraction image pixel value was determined by the difference between the most frequent colors in B and CE mode, that is, $\text{Mode}_{\text{CE-mode}} - \text{Mode}_{\text{B-mode}}$. Because the subtraction values range from -255 to 255 , they were normalized to a range of $0-255$ for imaging. This process was performed for each point by dividing the original images into segments, each with a width of 80 and a height of 120, resulting in a final subtraction image size of 80 pixels in width and 120 pixels in height. Labeling of the tumors in the subtraction images using the labeled areas in the CE mode and overlaying them on the subtraction images. The labeled subtraction images were used as training data, and the subtraction model (SM) was constructed in the same manner as the BRM (Figure 3c).

2.2.5 | Creation of combination model (CM) based on results from BRM and SM

To further improve the tumor recognition performance, a new model was constructed by learning the results of the two models (Figure 4). First, output images were generated from the training images for each of the two models (BRM and SM). Subsequently, based on the predicted

masked areas of tumors obtained from the two models, a model was developed to assess the reliability of the outcomes. The detailed algorithm for combining the BRM and SM is presented in Supplementary Method 2—Data S1. For training, 140 atypical and 35 tumor-free images were used.

2.3 | Evaluation of the three models

We assessed the performance of the three models for tumor detection. To achieve this, atypical tumor and tumor-free images were extracted from images of patients that were different from those used in the training data. First, these images were input into the BRM and SM for prediction, and the results were then used as input data for the CM to make decisions. Each model categorized the input data as either “tumor present” or “no tumor,” and these classifications were compared with the ground truth data. We quantitatively compared the performance to determine whether the tumor was accurately identified in this flow (Figure 4).

2.4 | Statistical analysis

The performance of the three models was evaluated using indices of accuracy, precision, recall, and F1-score. The recall was $TP / (TP + FN)$ (TP: true positive; FN: false negative). The F1-score is a metric that represents the harmonic

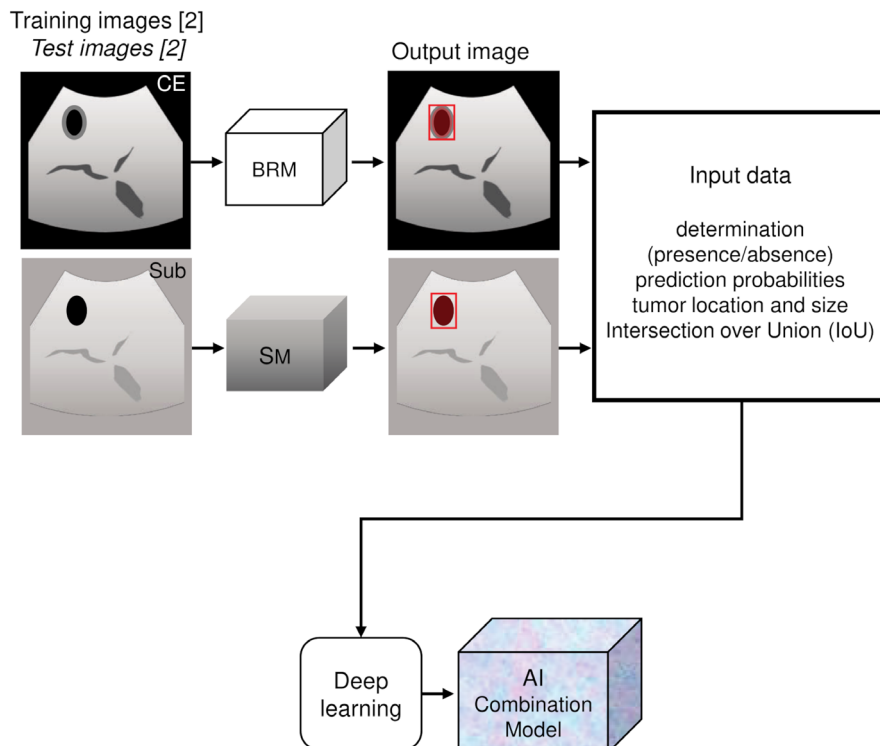


FIGURE 4 Constructing and implementing CM. The flow of construction and implementation of the combination model is shown in the schema, and the output images were generated from the training images for each BRM and SM. A combination model was developed using machine learning to evaluate the reliability of the outcomes using the results of BRM and SM. BRM, basic recognition model; CM, combination model; SM, subtraction model.

mean of precision and recall, providing a balanced evaluation of performance. This is beneficial for addressing class imbalances and is more reliable than accuracy because it considers both false positives and false negatives.

We analyzed the area under the receiver operating characteristic curve (AUROC) for each model. The accuracy, precision, recall, and F1-score were calculated using the optimum cutoff threshold derived from the area under the curve (AUC) analysis. The implementation was performed using Python 3.11,¹⁷ and the evaluation calculations were performed using scikit learn 1.2.^{18,19}

3 | RESULTS

A total of 121 patients were suitable for our study (Figure S1), and 2659 images containing tumors were extracted and labeled. We obtained 1786 typical images and 873 atypical images. Additionally, 478 images without tumors were extracted and used. To create BRM and SM, 1270 images were randomly selected from typical images and used as training data (training images[1]). Subsequently, 90 images each from atypical tumors and tumor-free images were randomly extracted and used as test data (test images[1]). To compare the three models, 175 images were used as training data (training images[2]) and 180 images as test data (test images[2]). Training image contained 140 atypical tumor images and 35 tumor-free images. Test images[2] contained 90 atypical tumor and 90 tumor-free images.

We identified the epochs wherein each model achieved its maximum AUC and then calculated the optimal thresholds within these epochs to ensure the best performance. The BRM was trained for 50 epochs, with the best-performing epoch selected from epoch 37. Similarly, the SM was trained for 50 epochs, and the model from epoch 9 was selected. The CM was trained for 50 epochs, and the model from epoch 41 was selected (Figure S2). Accuracy and other indices were compared using each optimal threshold.

Figure 5a,b show the predicted masked regions of the BRM and SM on the actual ultrasound images. The masked region was superimposed onto the tumor image. Tumor recognition was good in both models on typical images, with accuracies of 89.4% for BRM and 86.6% for SM (Table 1). In some cases, the BRM misidentified the vessels as tumors (Figure 5c, upper row). In contrast, vessels in the same locations were no longer identified by SM, and the true tumors were recognized (Figure 5c, lower row). However, even in cases where the tumor could be recognized by BRM (Figure 5d, upper row), SM sometimes failed to recognize tumors because of poor image generation of the liver parenchyma caused by ultrasound

attenuation and varying levels of contrast enhancement (Figure 5d, lower row).

In CM, we combined these two models to create a new tumor recognition model; tumors that were difficult to recognize by either BRM or CM could be detected by CM (Figure 5c,d, right). The accuracy of CM was 96.5%, which was higher than those of BRM and CM. The AUROC for BRM, SM, and CM are shown in Figure 6a: AUROC 0.94 for BRM, Figure 6b: 0.92 for SM, and Figure 6c: 0.99 for CM.

Three models were applied to the intraoperative ultrasonography video to display the masked areas (Figure S). The red masked areas indicate the areas recognized as tumors by BRM, green by SM, and blue by CM. In the BRM (red), the Glissonian sheath in the deep area was misidentified as a tumor. The tumor area is widely masked in the SM (green). In contrast, in the CM (blue), the true tumor area was masked without excess or deficiency.

4 | DISCUSSION

This study aimed to develop a system for the automatic detection of CLM using CE-IIOUS. We developed three different models (BRM, SM, and CM) that accurately detected CLM and displayed mask regions overlapping the tumors. Our method is novel in two respects. First, we used an algorithm to create a subtraction image by focusing on the difference in pixel values between the tumor and vessels prior to machine learning, to easily differentiate between the tumor and vessels. There are no previous reports incorporating the act of comparing pixel values on the ultrasound screen (something examiners do unconsciously) into an algorithm, and we successfully identified tumors that image learning alone could not detect. Second, we improved accuracy by integrating two tumor recognition models with different characteristics using machine learning. We found that models based solely on image learning and those incorporating algorithms exhibited different strengths and weaknesses in tumor recognition. Therefore, by combining these two models and further training them, we successfully constructed a model with a performance surpassing that of the individual models.

Computed tomography (CT), magnetic resonance imaging (MRI), and ultrasonography (US) are commonly used in clinical practice to screen and diagnose liver tumors. Advances in preoperative imaging modalities and surgical techniques have improved CLM resection rates, with 5-year survival rates after liver resection reported as 32%–58%.^{20–22} Since the only treatment with the best chance of cure for CLM is complete resection of the lesions,²³ accurate pre- and intraoperative identification of the number and location of lesions is essential. The accuracy of

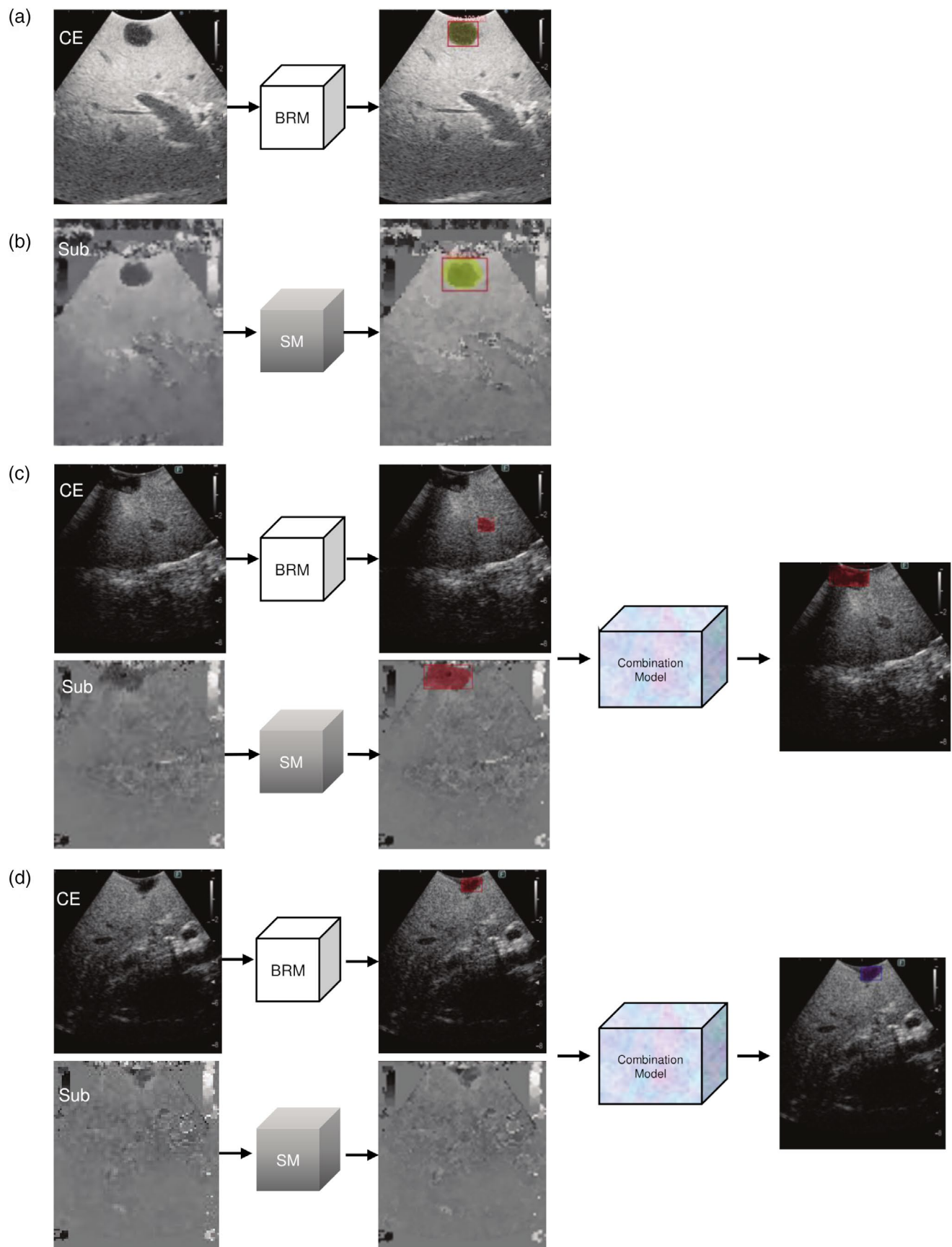


FIGURE 5 The typical images showing the predicted masked regions of BRM and SM. The regions predicted by BRM (a) and SM (b) were overlaid onto the tumor images. In small CLM, BRM sometimes misidentifies tumors while SM correctly identifies them (c). SM sometimes fails to recognize tumors owing to poor image generation, whereas BRM can detect them (d). In both cases, CM successfully detected tumors that were challenging to detect using either BRM or SM alone (c, d, right). BRM, basic recognition model; CM, combination model; SM, subtraction model.

TABLE 1 Comparison of three models.

	Accuracy	Precision	Recall	F1 score	Optimal threshold	AUROC
BRM	0.894	0.912	0.922	0.917	0.700	0.94
SM	0.866	0.882	0.911	0.896	0.638	0.92
CM	0.965	0.989	0.956	0.972	0.320	0.99

Abbreviations: AUROC, area under the receiver operating characteristic curve; BRM, basic recognition model; CM, combination model; SM, subtraction model.

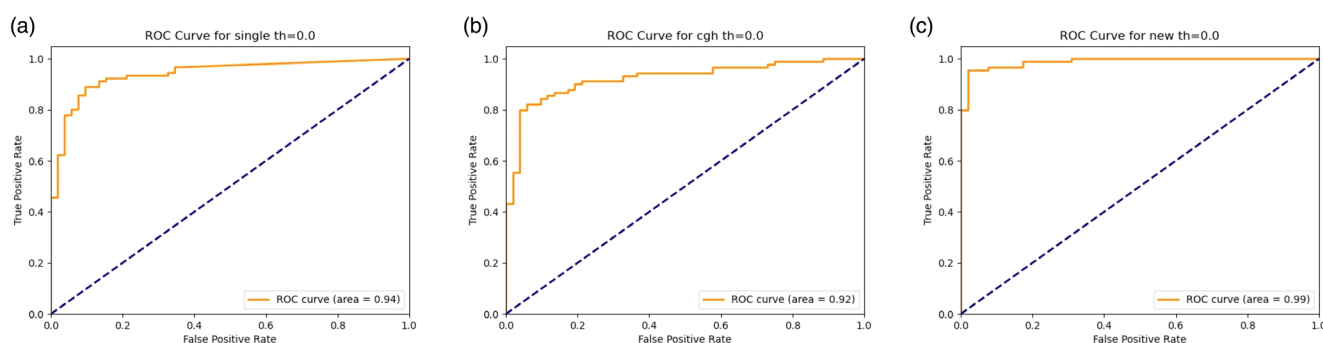


FIGURE 6 Comparison of the areas under the receiver operating characteristic curve. (a) BRM (b) SM and (c) CM. BRM, basic recognition model; CM, combination model; ROC, receiver operating characteristic curve; SM, subtraction model.

CE-IOUS (96%–97%) surpasses that of CT (72%–81%) and is comparable to that of MRI (73%–97%),^{24–26} making it crucial for real-time identification during surgery as it can be used intraoperatively.¹ However, CEUS depends on the examiner's skill and experience such as extracorporeal US,² and is even more challenging, as it requires comparing two screens simultaneously, which can lead to variability and challenges for less experienced practitioners.

In this study, we utilized artificial intelligence (AI) to support the detection of CLM in CE-IOUS and developed an automatic tumor recognition system that does not rely on the examiner's expertise, with an accuracy of 96.5% and precision of 98.9%, comparable to MRI or CE-IOUS interpreted by experts. Our system could improve the intraoperative CLM detection rates, promote complete resection of metastases, and increase the chances of cure, as it can deliver consistently high accuracy regardless of who performs the examination. Furthermore, when this system is introduced into clinical practice, no preoperative preparation is required. When a laptop computer equipped with this system (application) is connected to the ultrasound system, real-time analysis is performed, and the processed images are displayed on the monitor. The system can also

be set to sound an alert when a tumor is detected, allowing the examiner to make a diagnosis while viewing the monitor displaying the ultrasound images.

CEUS is useful for real-time spatial resolution and can detect smaller CLM lesions than can IOUS.²⁷ In fact, the diagnostic ability of CT and MRI decreases for smaller lesions, but CE-IOUS can detect lesions <5 mm, making it the final imaging modality for determining whether tumors can be resected.²⁸ However, even with CEUS, the detection of small CLM is sometimes difficult because of their atypical US imaging features,²⁹ making it difficult to differentiate them from vessels. Especially, the Glissonian sheath can appear as a circular hypoechoic area with a surrounding hyperechoic area when scanned transversely, potentially being mistaken for small hypoechoic tumors.

We developed an algorithm focusing on the difference in pixel values of the CLM between the B and CE mode. The echogenicity of the tumors and Glisson's sheath undergoes inverse changes before and after perflubutane injection. This inverse echogenicity change was used to create a subtraction image by subtracting the pixel values from the CE mode to the B mode; these data were

used to build the SM. The developed SM showed good accuracy (86.6%) and was effective in distinguishing between small tumors and vessel cross sections. The ability to quantify changes in echogenicity (pixel values) helped to objectively identify tumors within the ultrasound images. The method used to subtract the pixel values and incorporate them into the algorithm has not yet been used in the field of CE-IIOUS. Although this concept is simple, it is effective in differentiating between tumors and vessels. In practice, SM can exclude vessels and detect tumors in images where BRM misidentifies transverse images of vessels as lesions, making it difficult to detect small lesions. However, there are a few conditions where generating a suitable subtraction image is challenging due to liver parenchymal contrast or ultrasound attenuation, making tumor detection with SM not always straightforward. Considering these factors, we attempted to build a highly accurate model by learning from the prediction results of each model, which resulted in constructing the CM. The constructed CM had better accuracy (96.5%) than the two models alone, demonstrating that integrating the two models allowed us to build a highly accurate system. Machine learning with AI is a useful tool. However, it has the drawback of being a black box, making it difficult to identify the causes of unintended results. To overcome this problem, machine learning with AI must identify specific features of the desired outcome and guide the learning process using algorithms. Additionally, the approach of constructing a system in multiple directions and combining them for the final discrimination is effective for improving the quality of discrimination systems using machine learning.

Recently, AI applications in the medical field have been remarkable, spanning various areas. Many technologies have been commercialized and used as support systems. While US tends to rely on the examiner's skill and experience, AI suggests that screening accuracy can be improved by reducing human errors.³⁰ Although the field of liver US is not commercialized, numerous models exist to diagnose liver tumors.^{5,6} A previous study classified liver lesions using AI in IIOUS, including CLM, with 74.6% accuracy and 80.2% AUC. However, for lesions <3 cm, the accuracy dropped to 65.2%, and the AUC dropped to 67.6%, indicating challenges in evaluating small lesions.⁷

Previous studies on US images and AI primarily focused on the diagnosis of detected lesions. However, there is little research regarding accurate tumor detection using US. Particularly in CLM, real-time lesion detection during surgery is important to achieve complete and exact resection. To our knowledge, there are no reports evaluating the accuracy of automatic detection using IIOUS, including small lesions. We developed a system to accurately detect

and display small CLM lesions during IIOUS, including small lesions <2 cm. The system achieved detection accuracy of 96.5% and detection precision of 98.9% with an AUC of 0.99 for lesions including under 2 cm. Basically, the diagnosis of CLM is possible using B- and CE-mode US. However, accurate diagnosis without missing any tumors requires skill and experience. In addition, in difficult cases such as multiple lesions, small lesions, or proximity to vascular vessels, accurate diagnosis may be difficult and time consuming even for skilled operators. Our system can diagnose such difficult cases with an accuracy of 96.5%, which is close to that of experts, and is useful for supporting diagnosis and reducing screening time. It also has advantages when used by novice users who are not familiar with CLM diagnosis using CE-IIOUS.

This study had some limitations. First, as a characteristic issue of AI, it cannot explain why a tumor is detected, lacking the ability to provide a rationale for its judgments.

Second, as this study was conducted at a single institution, there was a potential bias in that CE-IIOUS images were collected by surgeons from a single facility. Our method required ultrasonographic images with good tumor visualization. Tumors are difficult to detect in cases of deep locations within the liver or fatty liver because of significant ultrasound attenuation in the background liver. Further studies are required to detect tumors in such cases.

Third, we did not specifically analyze the differences in accuracy based on tumor localization such as the liver segment, deep and surface localization, and proximity, and remoteness to blood vessels. Instead, we analyzed images that included a variety of tumor locations to build a comprehensive diagnostic system. Future studies would be required to further evaluate the impact of tumor localization on system performance. Additionally, we did not specifically compare the differences in surgical approaches including open, laparoscopic, and robot-assisted resection in this study. However, the analyzed images included all types of surgical approaches and were used equally in both training and test data sets. As long as B- and CE-mode images can be acquired through IIOUS, our system would be theoretically applicable regardless of the surgical approach. Future studies would be needed to evaluate the difference in surgical approaches.

Finally, this study focused on the identification of CLM, and thus, it is unclear whether it can be applied to other intrahepatic neoplastic lesions. Our system is designed to effectively detect tumors with contrast differences between B- and CE-mode images. From this perspective, since ICC and HCC have been reported to exhibit similar enhancement patterns to CLM, our system potentially has utility in differentiating these lesions. However, our system may not be effective for liver

abscess because abscess show different enhancement patterns from CLM. Future studies would be needed to evaluate the efficacy of our system for other intrahepatic lesions.

In conclusion, we successfully constructed a model to automatically detect CLM, including small lesions, with high precision using CE-IOUS. By integrating a model trained on CE-IOUS images containing CLM and a model constructed through image processing using pixel-value algorithms, we achieved improved detection accuracy.

ACKNOWLEDGMENTS

We would like to thank Editage (www.editage.jp) for English language editing.

FUNDING INFORMATION

This work was supported by Japan Society for the Promotion of Science KAKENHI (20K20214 to YM); and a grants-in-aid of the 106th annual congress of Japan Surgical Society Memorial Surgical Research Fund, Tokyo, Japan to Y.M.

CONFLICT OF INTEREST STATEMENT

The authors confirm there are no conflicts of interest.

ORCID

Nobuhisa Akamatsu  <https://orcid.org/0000-0003-1603-5147>

Yoshikuni Kawaguchi  <https://orcid.org/0000-0003-2986-3224>

Kiyoshi Hasegawa  <https://orcid.org/0000-0001-8734-740X>

REFERENCES

- Nielke MC, Bipat S, Stoker J. Diagnostic imaging of colorectal liver metastases with CT, MR imaging, FDG PET, and/or FDG PET/CT: a meta-analysis of prospective studies including patients who have not previously undergone treatment. *Radiology*. 2010;257(3):674–84.
- Reizine E, Mulé S, Luciani A. Focal benign liver lesions and their diagnostic pitfalls. *Radiol Clin North Am*. 2022;60(5):755–73.
- Ishikawa Y, Sugino T, Okubo K, Nakajima Y. Detecting the location of lung cancer on thoroscopic images using deep convolutional neural networks. *Surg Today*. 2023;53(12):1380–7.
- Katayama A, Aoki Y, Watanabe Y, Horiguchi J, Rakha EA, Oyama T. Current status and prospects of artificial intelligence in breast cancer pathology: convolutional neural networks to prospective vision transformers. *Int J Clin Oncol*. 2024;29:1648–68.
- Hu HT, Wang W, Chen LD, Ruan SM, Chen SL, Li X, et al. Artificial intelligence assists identifying malignant versus benign liver lesions using contrast-enhanced ultrasound. *J Gastroenterol Hepatol*. 2021;36(10):2875–83.
- Li W, Lv XZ, Zheng X, Ruan SM, Hu HT, Chen LD, et al. Machine learning-based ultrasonomics improves the diagnostic performance in differentiating focal nodular hyperplasia and atypical hepatocellular carcinoma. *Front Oncol*. 2021;11:544979.
- Barash Y, Klang E, Lux A, Konen E, Horesh N, Pery R, et al. Artificial intelligence for identification of focal lesions in intraoperative liver ultrasonography. *Langenbeck's Arch Surg*. 2022;407(8):3553–60.
- He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. *IEEE Trans Pattern Anal Mach Intell*. 2020;42(2):386–97.
- Lin TY, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. *IEEE Trans Pattern Anal Mach Intell*. 2020;42(2):318–27.
- Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell*. 2017;39(6):1137–49.
- Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, et al. Automatic differentiation in pytorch. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.
- The Linux Foundation. PyTorch 2.1. Available from: <https://pytorch.org/>. Accessed 12 Nov 2024.
- Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, et al. Microsoft coco: common objects in context. *Computer vision—ECCV 2014. 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13*: Springer; 2014.
- COCO-Common Objects in Context Available from: <https://cocodataset.org/#home>. Accessed 12 Nov 2024.
- The Linux Foundation. PyTorch-MODEL and Pre-Trained Weights. Available from: <https://pytorch.org/vision/stable/models.html#instance-segmentation> Accessed 12 Nov 2024.
- Diederik P, Kingma JB. Adam: A Method for Stochastic Optimization. *International Conference for Learning Representations*; San Diego 2015.
- Python Software Foundation. Python 3.11. Available from: <https://www.python.org/downloads/release/python-3110/>. Accessed 12 Nov 2024.
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. *The Journal of Machine Learning Research*. 2011;12:2825–30.
- scikit learn 1.2 Available from: <https://scikit-learn.org/1.2/>. Accessed 12 Nov 2024.
- Abdalla EK, Vauthey JN, Ellis LM, Ellis V, Pollock R, Broglio KR, et al. Recurrence and outcomes following hepatic resection, radiofrequency ablation, and combined resection/ablation for colorectal liver metastases. *Ann Surg*. 2004;239(6):818–27.
- Nordlinger B, Sorbye H, Glimelius B, Poston GJ, Schlag PM, Rougier P, et al. Perioperative FOLFOX4 chemotherapy and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC 40983): long-term results of a randomised, controlled, phase 3 trial. *Lancet Oncol*. 2013;14(12):1208–15.
- Kato T, Yasui K, Hirai T, Kanemitsu Y, Mori T, Sugihara K, et al. Therapeutic results for hepatic metastasis of colorectal cancer with special reference to effectiveness of hepatectomy: analysis of prognostic factors for 763 cases recorded at 18 institutions. *Dis Colon Rectum*. 2003;46(10 Suppl):S22–S31.
- Minagawa M, Makuuchi M, Torzilli G, Takayama T, Kawasaki S, Kosuge T, et al. Extension of the frontiers of surgical indications in the treatment of liver metastases from colorectal cancer: long-term results. *Ann Surg*. 2000;231(4):487–99.
- Leen E, Ceccotti P, Moug SJ, Glen P, MacQuarrie J, Angerson WJ, et al. Potential value of contrast-enhanced intraoperative

- ultrasonography during partial hepatectomy for metastases: an essential investigation before resection? *Ann Surg.* 2006;243(2):236–40.
25. Arita J, Ono Y, Takahashi M, Inoue Y, Takahashi Y, Matsueda K, et al. Routine preoperative liver-specific magnetic resonance imaging does not exclude the necessity of contrast-enhanced intraoperative ultrasound in hepatic resection for colorectal liver metastasis. *Ann Surg.* 2015;262(6):1086–91.
 26. Berger-Kulemann V, Schima W, Baroud S, Koelblinger C, Kaczirek K, Gruenberger T, et al. Gadoxetic acid-enhanced 3.0 T MR imaging versus multidetector-row CT in the detection of colorectal metastases in fatty liver using intraoperative ultrasound and histopathology as a standard of reference. *Eur J Surg Oncol.* 2012;38(8):670–6.
 27. Desolneux G, Isambert M, Mathoulin-Pelissier S, Dupré A, Rivoire M, Cattena V, et al. Contrast-enhanced intra-operative ultrasound as a clinical decision making tool during surgery for colorectal liver metastases: the ULIIS study. *Eur J Surg Oncol.* 2019;45(7):1212–8.
 28. Arita J, Ono Y, Takahashi M, Inoue Y, Takahashi Y, Saiura A. Usefulness of contrast-enhanced intraoperative ultrasound in identifying disappearing liver metastases from colorectal carcinoma after chemotherapy. *Ann Surg Oncol.* 2014;21(Suppl 3):S390–S397.
 29. Lim GH, Koh DC, Cheong WK, Wong KS, Tsang CB. Natural history of small, "indeterminate" hepatic lesions in patients with colorectal cancer. *Dis Colon Rectum.* 2009;52(8):1487–91.
 30. Nishida N, Kudo M. Artificial intelligence in medical imaging and its application in sonography for the management of liver tumor. *Front Oncol.* 2020;10:594580.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Takayama M, Ito K, Karako K, Mihara Y, Sasaki S, Ichida A, et al. An artificial intelligence-based recognition model of colorectal liver metastases in intraoperative ultrasonography with improved accuracy through algorithm integration. *J Hepatobiliary Pancreat Sci.* 2025;32:58–68. <https://doi.org/10.1002/jhbp.12089>