Original Article

Potential of automatic diagnosis system with linked color imaging for diagnosis of *Helicobacter pylori* infection

Takeshi Yasuda,[1] Tomoyuki Hiroyasu,[2] Satoru Hiwa,[2] Yuto Okada,[3] Sadanari Hayashi,[1] Yuki Nakahata,[1] Yuriko Yasuda,[1] Tatsushi Omatsu,[1] Akihiro Obora,[1] Takao Kojima,[1] Hiroshi Ichikawa[2] and Nobuaki Yagi[1]

[1]Department of Gastroenterology, Asahi University Hospital, Gifu, [2]Faculty of Life and Medical Sciences, and [3]Graduate School of Life and Medical Sciences, Doshisha University, Kyoto, Japan

Background and Aim: It is necessary to establish universal methods for endoscopic diagnosis of *Helicobacter pylori* (HP) infection, such as computer-aided diagnosis. In the present study, we propose a multistage diagnosis algorithm for HP infection.

Methods: The aims of this study are to: (i) to construct an interpretable automatic diagnostic system using a support vector machine for HP infection; and (ii) to compare the diagnosis capability of our artificial intelligence (AI) system with that of endoscopists. Presence of an HP infection determined through linked color imaging (LCI) was learned through machine learning. Trained classifiers automatically diagnosed HP-positive and -negative patients examined using LCI. We retrospectively analyzed the new images from 105 consecutive patients; 42 were HP positive, 46 were post-eradication, and 17 were uninfected. Five endoscopic images per case taken from different areas were read into the AI system, and used in the HP diagnosis.

Results: Accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of the diagnosis of HP infection using the AI system were 87.6%, 90.4%, 85.7%, 80.9%, and 93.1%, respectively. Accuracy of the AI system was higher than that of an inexperienced doctor, but there was no significant difference between the diagnosis of experienced physicians and the AI system.

Conclusions: The AI system can diagnose an HP infection with significant accuracy. There remains room for improvement, particularly for the diagnosis of post-eradication patients. By learning more images and considering a diagnosis algorithm for post-eradication patients, our new AI system will provide diagnostic support, particularly to inexperienced physicians.

Key words: artificial intelligence, Helicobacter pylori infection diagnosis, linked color imaging, machine learning, support vector machine

INTRODUCTION

CONTINUOUS INFECTION WITH *Helicobacter pylori* (HP) has been reported to be one of the biggest factors in gastric cancer.[1,2] To diagnose or classify chronic gastritis by endoscopy, the Kimura-Takemoto classification,[3] the updated Sydney system,[4] and the Kyoto classification[5,6] are widely used. Lately, a new image-enhanced endoscopy system, called linked color imaging (LCI), was developed by Fujifilm Co. (Tokyo, Japan). A characteristic of LCI is enhancement of the slight difference in mucosal reddish color.

Some recent studies have reported that LCI is significantly useful for the diagnosis of active HP infection or gastric metaplasia.[7,8] However, the endoscopic diagnosis of HP infection does not have objective indicators, and the diagnostic ability differs between doctors. Therefore, it is necessary to establish interpretable universal methods for an endoscopic diagnosis of an HP infection, such as a computer-aided diagnosis (CAD). The aims of the present study are to: (i) construct an interpretable automatic diagnostic system using a support vector machine (SVM) for HP infection; and (ii) to compare the diagnosis capability of our AI system with that of endoscopists.

In recent years, AI using deep learning (especially for convolutional neural networks; CNN) has been applied to gastroenterology.[9–15] However, CNN have a disadvantage in that it is difficult to investigate the rationale for a diagnosis

(e.g. understanding the discriminative features for the diagnosis) from the learned models because they construct and optimize the nonlinear input-output mapping through end-to-end supervised learning. It is very important not only to improve the accuracy of diagnosis but also to expand the knowledge of endoscopists by understanding which features of the images are discriminative between the HP-positive or -negative samples. Furthermore, if the expert endoscopists already have empirical knowledge for the diagnosis, it is also possible to improve the accuracy by embedding the prior knowledge into the machine-learning model. In this case, the conventional machine-learning process is effective.

In the present study, we constructed a machine-learning-based algorithm for an automatic diagnosis system for HP infection using LCI images. We used a SVM as a classification algorithm and evaluated its accuracy by comparing it with the diagnosis of the endoscopists.

METHODS

Method for creating the training data

FOR THE TRAINING data, we used four endoscopic images each from 32 cases, which were taken from the lesser (angle-lower body and middle-upper body) and greater (angle-lower body and middle-upper body) curvature. All 32 cases were observed by board-certified endoscopists and checked for HP infection based on more than two different examinations: a histological examination, a serum antibody test (baseline level: 0–10.0 U/mL), a stool antigen test (positive or negative), and/or a ^{13}C-urea breath test (baseline level: 0–2.4‰). Those images were observed at Asahi University Hospital from April 2015 to August 2015. Details of the 32 cases are as follows (Table 1). Nineteen cases were male, 13 cases were female. Regarding HP infection of the subjects, 14 cases were HP positive, 18 cases were HP uninfected. Regarding the atrophic gland border (Kimura-Takemoto classification), 24 cases were closed type and eight cases were open type.

Classification of LCI images using feature values

In our AI system, LCI images were classified into two types based on a slight difference in redness in high-hue images (red-purple: $0 \leqq H \leqq 45$) and low-hue images (apricot-red, $315 \leqq H < 360$; Fig. 1). We selected high hue images as the region of interest (ROI). To extract the ROI, some preprocessing was applied. First, a closing operation, which applies dilation followed by an erosion, was used to eliminate small holes in the image as noise. Saturation and value (S and V) of the Hue·Saturation·Value (HSV) color space were used to

Table 1 Baseline characteristics of patients recruited for the training data

	N = 32
Age (y), median (IQR)	59 (38–89)
Gender, n (%)	
Male	19
Female	13
HP infection	
Current infection	14
Uninfected	18
Mucosal atrophy	
C-I	11
C-II	7
C-III	6
O-I	1
O-II	6
O-III	1

HP, *Helicobacter pylori*.

determine the noise. For example, saturation (S) of the halation area was low, and saturation and value (S and V) of the endoscope fiber cable were also low. Therefore, an area of low saturation or low value was eliminated as noise (Fig. 2). Second, a raster scan to detect the connected pixels (component) which had a hue value greater than 128 ($H \geqq 128$) in the image as the mucosal area was applied (Fig. 3a,b), and then the connected components within 1400 pixels were eliminated as the non-ROI area (Fig. 3b,c).

Construction of diagnosis algorithm of HP infection using machine learning

Based on the ROI, endoscopic images were divided into two patterns (high- and low-hue images). Those images were calculated to determine the feature value of the mucosal color separately, and we constructed the SVM classifier using a radial basis function kernel. Among the 128 images, 86 were classified into low-hue, and 42 images were classified into high-hue images. The feature values used in the SVM were ratio of the ROI to the entire image size, average and median hue values in high-hue images (Fig. 4a), and mode of the saturation value and the median and variance of the hue value in low-hue images (Fig. 4b).

Then, the trained classifiers diagnosed HP infection of one image each automatically and made a comprehensive determination of the presence or absence of HP (Fig. 5).

Diagnosis of HP infection based on new endoscopic images using the AI system

To investigate the versatility of the newly constructed SVM, we prepared five unlearned endoscopic images per case,
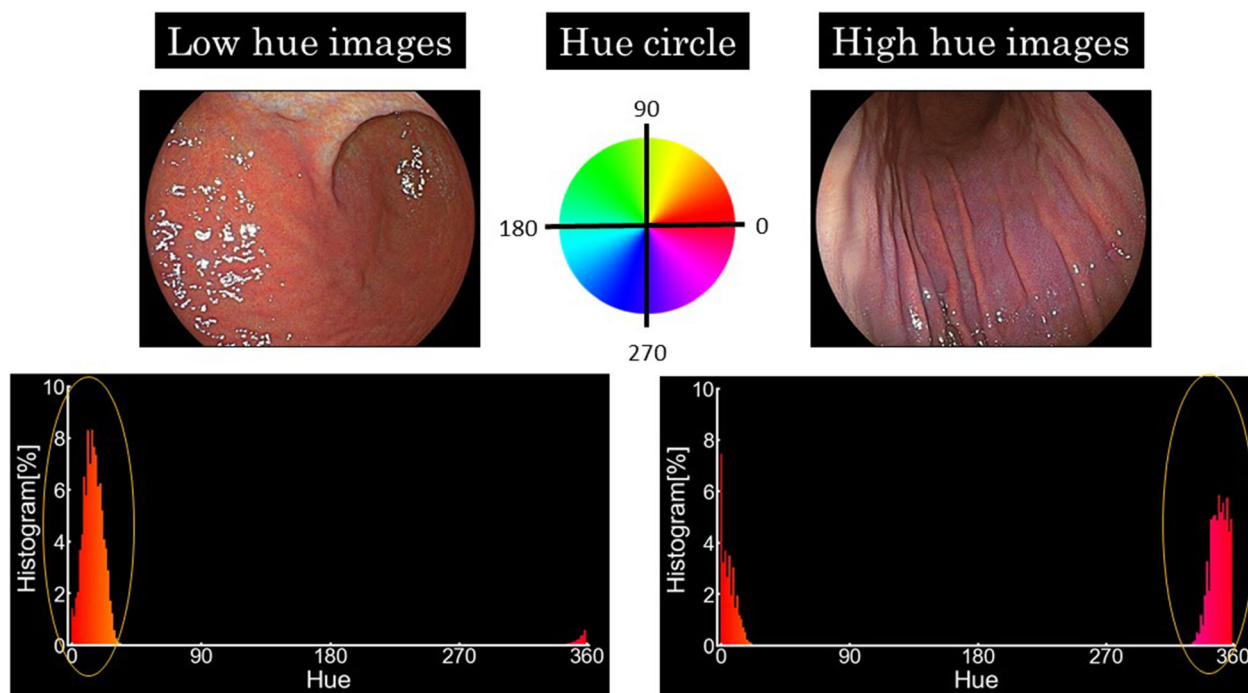
Figure 1 Linked color imaging images were classified into two groups based on a difference in the hue value in Hue·Saturation·Value color space which reflects a difference in mucosal redness: high-hue (red-purple: $0 \leqq H \leqq 45$) and low-hue images (apricot-red, $315 \leqq H < 360$).

which were taken from the lesser (angle-lower body and middle-upper body) and greater (angle-lower body and middle-upper body) curvature and the fornix. These five LCI images per patient were read into the AI system, and final judgment regarding HP-positive/negative was made. When the diagnoses of an image were not the same, the final diagnosis was made by majority rule. Three doctors also evaluated the same five LCI images. The three doctors were as follows: A, an expert involved in the development of LCI; B, a gastroenterology specialist; and C, a senior resident. The doctors used the Kyoto classification for diagnosing HP infection.

Patients

We retrospectively analyzed the images from 105 consecutive patients who were checked for HP infection based on more than two different examinations that used the same criteria as the training data. The images were taken using LCI mode at Asahi University Hospital from January 2017 to January 2018. All images were taken by board-certified of endoscopists. Characteristics of the patients were as follows: (i) median age, 64 years old (26–88); (ii) gender, 61 male

and 44 female; (iii) HP infection, 42 HP positive, 46 post-eradication, and 17 uninfected; and (iv) mucosal atrophy, 46 closed type and 59 open type (Kimura-Takemoto classification; Table 2). This study included 46 cases of post-HP-eradicated gastric mucosa. The patients underwent eradication therapy in 2006–2017. More than 1 year (average of 5.6 years) had passed since HP was successfully eradicated after undergoing endoscopy. We excluded patients who were not checked for HP infection using more than two examinations or who showed mucosal atrophy but did not have a history of HP eradication and who did not have images taken from all five perspectives in the LCI mode (Fig. 6).

Endoscopy system

All examinations were carried out using a LASEREO system (VP-4450 or -7000, and LL-4450 or -7000) and an upper gastrointestinal endoscope (EG-L590ZW or EG-L600ZW) (Fujifilm Co.). These endoscopic systems have both LCI and WLI modes. Color emphasis and structural emphasis of the LCI mode were set to C1 and B5, respectively.
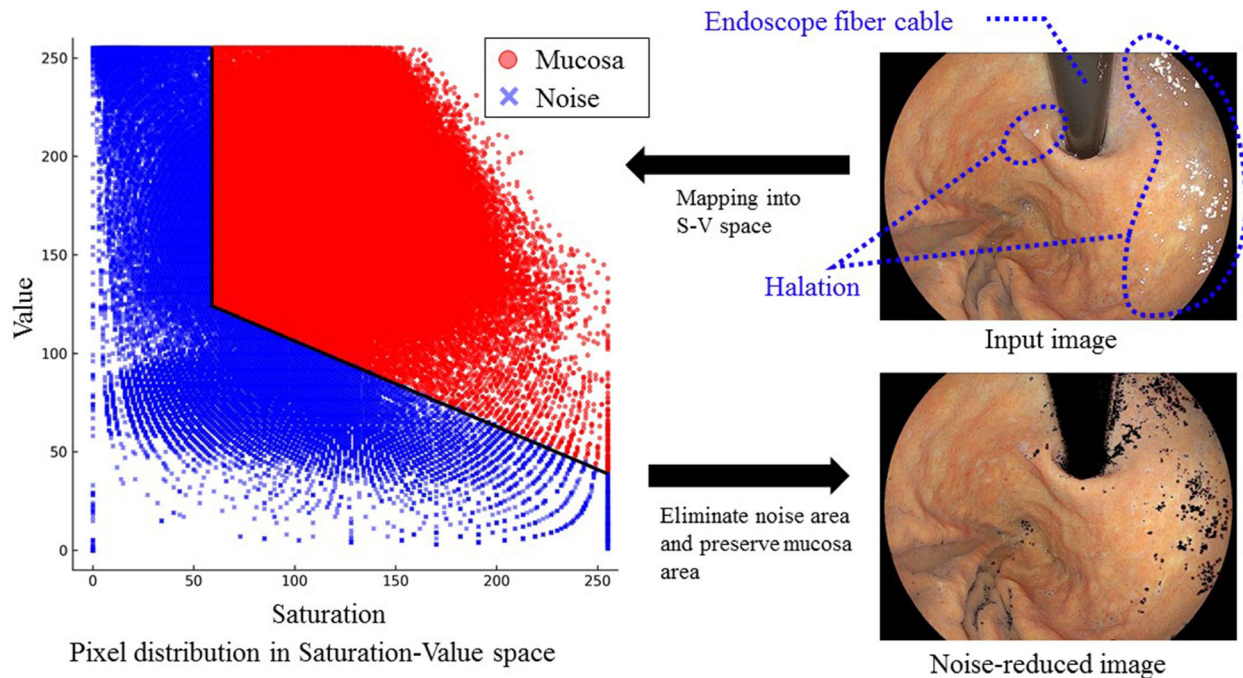
Figure 2 Schematic illustration of the preprocessing used in the present study to remove noise in endoscopic images. Pixels with low saturation or low value in Hue·Saturation·Value color space were extracted and eliminated as noise representing the halation and the endoscope fiber cable. S-V space, saturation-value space.
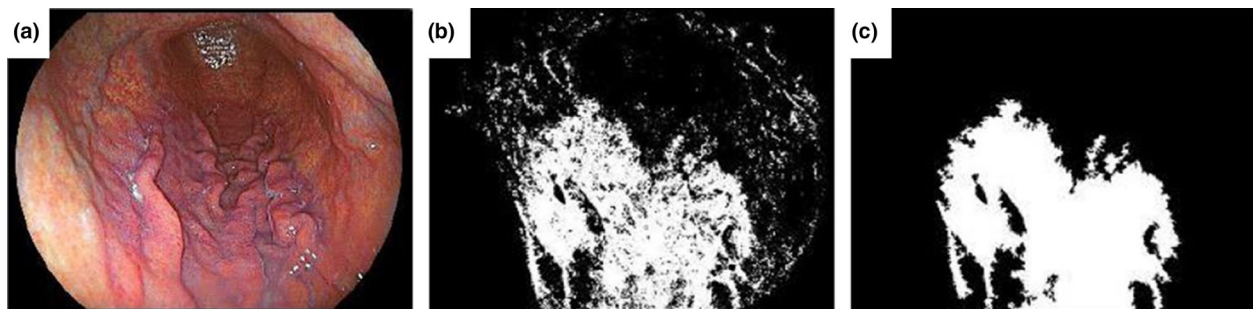

Figure 3 Labeling process is applied to detect connected pixels (component) which have a hue value >128 (H > 128) in the images as the mucosal area through a raster scan (a→b). Then the connected components within 1400 pixels are eliminated (b→c) as the non-region of interest area.

RESULTS

FIVE HUNDRED AND twenty-five images of the 105 cases were divided into two groups automatically. The first group was the low-hue group with 490/525 images (182 images were of HP-positive cases and 308 images were of post-HP-eradication cases); the second group was the high-hue group with 35/525 images (28 images were of HP-positive cases and seven images were of post-HP-eradication cases).

Levels of accuracy, sensitivity, specificity, PPV, and NPV of diagnosis of HP infection using the AI system were 87.6% (92/105), 90.5% (38/42), 85.7% (54/63), 80.9% (38/47), and 93.1% (54/58), respectively. In contrast, those from a diagnosis by A were 90.5% (95/105), 92.9% (39/42), 88.9% (56/63), 84.8% (39/46), and 94.9% (56/59),
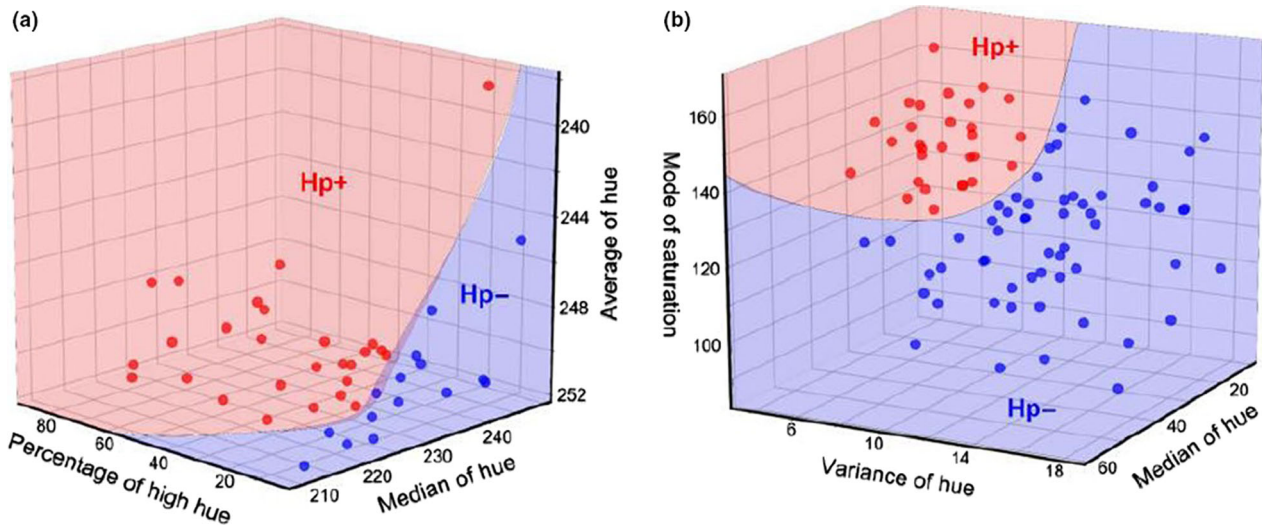
Figure 4 Feature values used in the support vector machine model for each of the high-hue and low-hue images. (a) Distribution of the pixels in the 3-D space formed by the percentage of the high-hue value (ratio of the region of interest to the entire image size), median hue value, and average hue value in high-hue images. (b) Distribution of pixels in the 3-D space formed by variance of hue value, median hue value, and mode of saturation value in low-hue images. Hp, *Helicobacter pylori*.

respectively. Those from a diagnosis by B were 89.5% (94/105), 85.7% (36/42), 92.1% (58/63), 87.8% (36/41), and 90.6% (58/64), respectively. Finally, those from a diagnosis by C were 86.7% (91/105), 92.9% (39/42), 82.5% (52/63), 78.0% (39/50), and 94.5% (52/55), respectively.

Accuracy of the AI system was higher than that of the inexperienced doctor (doctor C), but there was no significant difference between the diagnosis of the physicians and the AI system (between AI and doctor A, $P = 0.51$; between AI and doctor B, $P = 0.67$; between AI and doctor C, $P = 0.84$, chi-squared test).

Subanalysis of the patients divided with respect to state of HP infection

First, we conducted a subanalysis of the HP-positive cases ($n = 42$). Among the 42 cases, 27 were male and 15 were female. Mean age was 59.2 (26–88) years. Atrophic border was as follows: closed type, 14 cases; open type, 28 cases. Accuracies of the AI system, doctors A, B, and C were 90.5% (38/42), 92.9% (39/42), 85.7% (36/42), and 92.9% (39/42), respectively.

Second, we conducted a subanalysis of post-HP-eradication patients ($n = 46$). Among the 46 patients, 27 were male and 19 were female. Mean age was 69.5 (50–88) years. Atrophic border was as follows: closed type, 15 cases; open type, 31 cases. Accuracies of the AI system, doctors A, B, and C were 82.6% (38/46), 87.0% (40/46), 89.1% (41/46), and 76.1% (35/46), respectively.

Subanalysis of AI diagnosis for each image of stomach area

We conducted a subanalysis of 525 images of 105 cases. Based on the location, accuracies of the lesser curvature of the angle-lower body, lesser curvature of the middle-upper body, fornix, greater curvature of the angle-lower body, and greater curvature of the middle-upper body were 76.2% (80/105), 88.6% (93/105), 69.5% (73/105), 77.1% (81/105), and 73.3% (77/105), respectively (Table 3). Accuracy of the lesser curvature of the middle-upper body was significantly higher than that of the fornix and the greater curvature of the middle-upper body ($P < 0.01$, chi-squared test).

Subanalysis of AI diagnosis based on hue of each image

We conducted a subanalysis of the hue of each image. Among all images, 93.3% (490/525) were classified into low-hue images, whereas 6.7% (35/525) were classified into high-hue images.

Accuracies of the low-hue images of HP-uninfected, HP-positive patients, and post-HP-eradication patients were 89.4% (76/85), 73.1% (133/182), and 75.8% (169/223), respectively. Accuracy of the high-hue images of HP-positive patients was 100.0% (28/28), whereas that of post-HP-eradication patients was 0.0% (0/7) (Table 4).
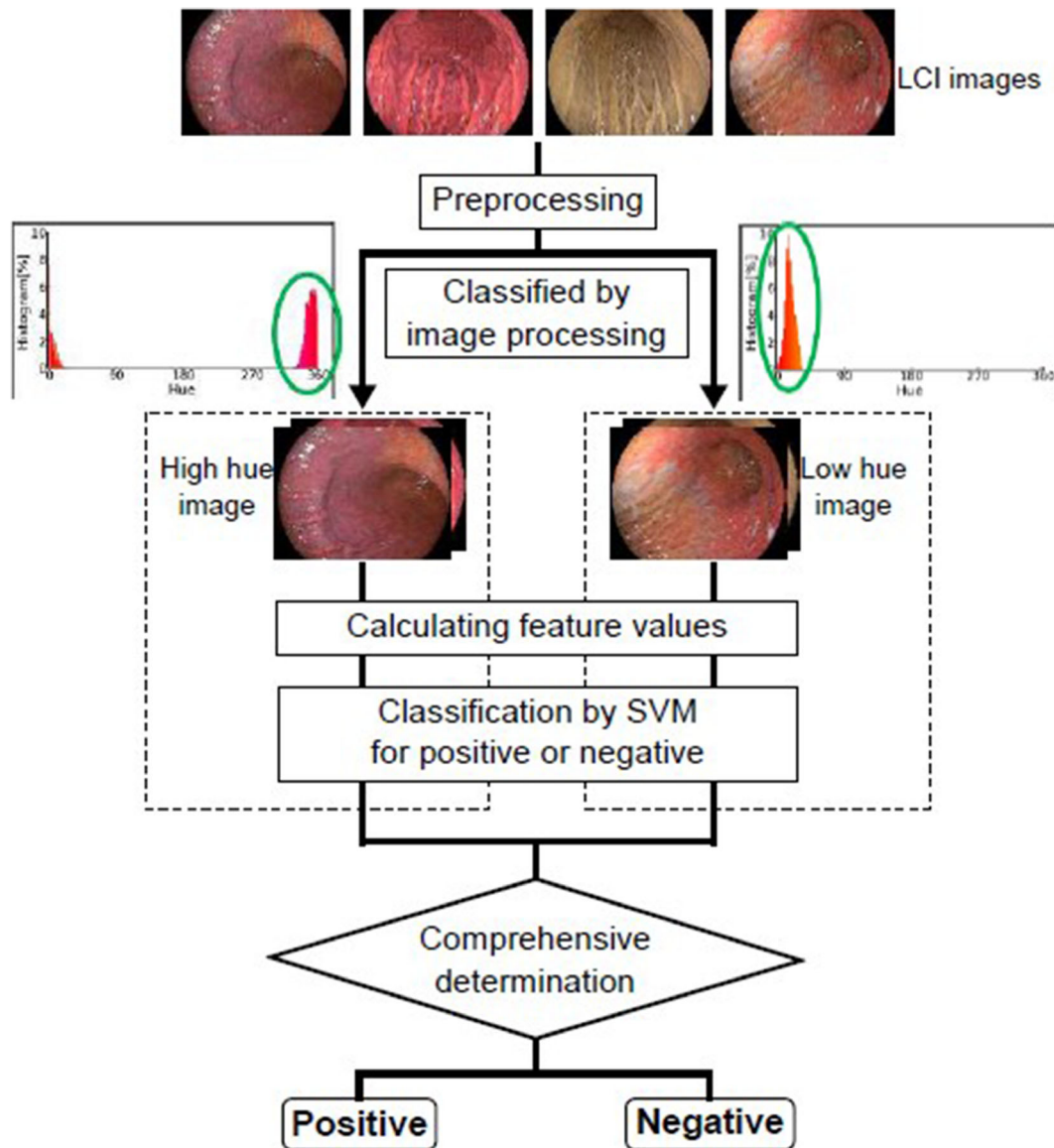
Figure 5 Schematic flow of the proposed machine-learning algorithm to diagnose *Helicobacter pylori* infection. LCI, linked color imaging; SVM, support vector machine.

DISCUSSION

IN THE PRESENT study, we combined LCI with a newly constructed AI system trained by SVM; the system uses an interpretable algorithm for the diagnosis of HP infection to support the clinical diagnosis of doctors.

Usefulness of LCI regarding the inflammation of gastrointestinal mucosa has been previously reported. Uchiyama *et al.*[16] showed that the mucosal redness enhanced by LCI was correlated with mucosal inflammation in ulcerative colitis. Dohi *et al.*[7] reported the usefulness of LCI for diagnosing HP-positive and post-HP-eradication cases by enhancing endoscopic images of diffuse redness. Takeda *et al.*[17] reported that LCI enhances color variation of the mucosal change caused by HP infection by checking against the Kyoto classification. We focused on these color variations as a feature value of our machine-learning system.

In recent years, the innovation of AI technology has been accelerating remarkably. In the field of gastroenterology, Mori *et al.*[18] reported a CAD system for the diagnosis of

Table 2 Baseline characteristics of patients recruited for the present study

	N = 105
Age (y), median (IQR)	64 (26–88)
Gender, *n* (%)	
Male	61
Female	44
HP infection	
Current infection	42
Post-eradication	46
Uninfected	17
Mucosal atrophy	
C-I	11
C-II	20
C-III	15
O-I	26
O-II	26
O-III	7

HP, *Helicobacter pylori*.

small colorectal polyps. Misawa *et al.*[19] reported a CAD system for the detection of colorectal polyps. With regard to upper digestive diseases, some studies have already reported the usefulness of CNN-based AI systems for diagnosing HP infection[9–13] and early detection of gastric neoplasms.[13–15] Shichijo *et al.*[9,10] reported a CNN-based AI system that diagnoses HP infections. Nakashima *et al.*[12] combined a CNN-based AI system with blue laser imaging and LCI. Their AI systems showed high accuracy; however, post-HP-eradication cases were excluded from their studies. As the importance of HP eradication is recognized worldwide, the number of post-HP-eradication cases is increasing. It is also

important to construct an algorithm for post-HP-eradication cases.

Although the use of CNN is one of the most popular deep-learning approaches for automatically extracting geometric features and differentiating between HP-positive and -negative images, CNN have a disadvantage in that it is difficult to investigate the rationale of their diagnosis. It is particularly important to explain the ground and process of the result obtained. To explore and extract the interpretable features of images that doctors acquire implicitly as empirical knowledge through their experience in the clinical field, we expect that the doctor's expertise can be embodied as a visible system with scientific evidence. Furthermore, this investigation would also be returned beneficially to endoscopists to expand their knowledge. Hence, we chose classical machine learning using SVM in the current study. Therefore, our proposed system should not be compared with deep-learning methodologies such as CNN-based classifiers which require preliminary determination of a feature-extraction process. The comparison between them would go beyond the scope of this study. Our new AI system classified the input images into high- and low-hue groups, and then extracted the ROI for classification. The SVM models were constructed for each group based on the feature values calculated within the ROI and the images of each group were separately classified into HP+ or HP− using the corresponding SVM model.

In the present study, we selected five endoscopic images that covered the entire gastric mucosa: lesser curvature (angle-lower and middle-upper body), greater curvature (angle-lower and middle-upper body), and fornix images. The accuracy of this AI system is superior to that of an
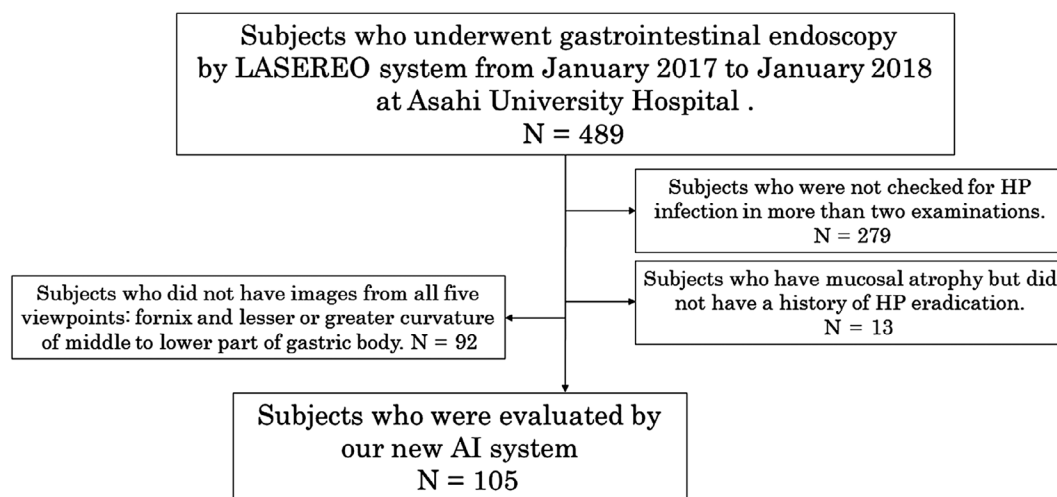

Figure 6 Flowchart of patients recruited in the present study. AI, artificial intelligence; HP, *Helicobacter pylori*.

Table 3 Subanalysis of the accuracy of AI diagnosis for each image based on location

Accuracy of AI diagnosis of each portion of the stomach

| | | Lesser curvature | | Fornix | Greater curvature | |
		Angle-lower body	Middle-upper body		Angle-lower body	Middle-upper body
Total		76.2% (80/105)	88.6% (93/105)	69.5% (73/105)	77.1% (81/105)	73.3% (77/105)
HP	Active	76.2% (32/42)	88.1% (37/42)	59.5% (25/42)	78.6% (33/42)	69.0% (29/42)
	After eradication	71.7% (33/46)	89.1% (41/46)	73.9% (34/46)	73.9% (34/46)	67.4% (31/46)
	Uninfected	88.2% (15/17)	88.2% (15/17)	82.4% (14/17)	82.4% (14/17)	100% (17/17)

AI, artificial intelligence; HP, *Helicobacter pylori*. [Correction added on 29 October 2019, after first online publication: Figures included in Table 3 have been adjusted to reflect the correct aspect ratio.]

Table 4 Subanalysis of the accuracy of AI diagnosis for each image based on hue

		Low-hue images	High-hue images
Distribution		93.3% (490/525)	6.7% (35/525)
Accuracy of AI diagnosis of each hue image			
Total		77.1% (378/490)	80% (28/35)
HP	Positive	73.1% (133/182)	100.0% (28/28)
	After eradication	75.8% (169/223)	0.0% (0/7)
	Uninfected	89.4% (76/85)	None

AI, artificial intelligence; HP, *Helicobacter pylori*.

inexperienced doctor but cannot surpass the diagnostic capability of an experienced doctor. Based on a subanalysis of each endoscopic image, the lesser curvature of the angle-lower body showed the greatest accuracy. In the diagnostic process of HP infection, further consideration regarding the area of the endoscopic images applied, or the number of images, is warranted for a complete diagnosis. Based on a subanalysis of each hue, the group of high-hue images can diagnose HP-positive patients with 100% accuracy, but could not detect post-eradication patients. This result indicates that images that have a high hue owing to map-like redness are difficult for our AI system to diagnose. The group of low-hue images can be used to diagnose HP-uninfected cases with high accuracy, whereas the accuracy for HP-positive or post-eradication patients is not sufficient for clinical adaptation. In the first training data set, there were no post-HP-eradication cases. We believe that the cases uninfected with HP could represent HP-negative (including post-eradicated) cases because the reddish color of LCI is correlated with mucosal inflammation. As a result, our system misdiagnosed the high-hue images with intestinal metaplasia and the low-hue images of HP-positive or post-eradication cases.

The present study has several limitations. First, it was conducted at a single center and the collected endoscopic images were few in number. Second, classification of the 105 unlearned cases showed a low ratio of high-hue images. Thus, we have to modulate the reference value when selecting high-hue images.

In a future study, we aim to include post-eradication cases in training data and investigate the feature value of intestinal metaplasia particularly "map-like redness", which was reported to show a lavender color in the LCI mode.[20]

Currently, using the proposed AI system in double-checking procedures in clinical practice is premature and this problem is not limited to our proposed system. In the future, it would be more suitable for developing countries, where the diagnosis and eradication of HP infection can be approved for patients with an endoscopic diagnosis of chronic gastritis.

CONCLUSION

THE AI SYSTEM was able to diagnose HP infections with significant accuracy. However, there is still room for improvement, particularly regarding the diagnostic capability for post-eradication patients. By learning more images and considering a diagnosis algorithm for post-eradication patients, our new AI system will be able to support doctors, particularly inexperienced physicians.

CONFLICTS OF INTEREST

AUTHORS DECLARE NO conflicts of interest for this article.

REFERENCES

1 Correa P. Human gastric carcinogenesis: a multistep and multifactorial process--First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention. *Cancer Res.* 1992; **52**: 6735–40.

2 Uemura N, Okamoto S, Yamamoto S *et al. Helicobacter pylori* infection and the development of gastric cancer. *N. Engl. J. Med.* 2001; **345**: 784–9.

3 Kimura K, Takemoto T. An endoscopic recognition of the atrophic border and its significance in chronic gastritis. *Endoscopy* 1969; **1**: 87–97.

4 Dixon MF, Genta RM, Yardley JH *et al*. Classification and grading of gastritis: the updated Sydney system. *Am. J. Surg. Pathol.* 1996; **20**: 1161–81.

5 Haruma K, Kato M, Inoue K *et al. Kyoto Classification of Gastritis*. Tokyo, Japan: Nihon Medical Center, 2017.

6 Sugano K, Tack J, Kuipers EJ *et al*. Kyoto global consensus report on *Helicobacter pylori* gastritis. *Gut* 2015; **64**: 1353–67.

7 Dohi O, Yagi N, Onozawa Y *et al*. Linked color imaging improves endoscopic diagnosis of active *Helicobacter pylori* infection. *Endosc. Int. Open* 2016; **04**: E800–5.

8 Osawa H, Miura Y, Takezawa T *et al*. Linked color imaging and blue laser imaging for upper gastrointestinal screening. *Clin. Endosc.* 2018; **51**: 513–26.

9 Shichijo S, Nomura S, Aoyama K *et al*. Application of convolutional neural networks in the diagnosis of *Helicobacter pylori* infection based on endoscopic images. *EBioMedicine* 2017; **25**: 106–11.

10 Shichijo S, Endo Y, Aoyama K *et al*. Application of convolutional networks for evaluating *Helicobacter pylori* infection status on the basis of endoscopic images. *Scand. J. Gastroenterol.* 2019; **54**: 158–63.

11 Itoh T, Kawahira H, Nakashima H, Yata N. Deep learning analyzed *Helicobacter pylori* infection by upper gastrointestinal endoscopy images. *Endosc. Int. Open* 2018; **6**: E139–44.

12 Nakashima H, Kawahira H, Kawachi H *et al*. Artificial intelligence diagnosis of *Helicobacter pylori* infection using blue laser imaging-bright and linked color imaging: a single-center prospective study. *Ann. Gastroenterol.* 2018; **31**: 462–8.

13 Mori Y, Kudo SE, Hussein E *et al*. Artificial intelligence and upper gastrointestinal endoscopy: current status and future perspective. *Dig. Endosc.* 2019; **31**: 378–88.

14 Wu L, Zhou W, Wan X *et al*. A deep neural network improves endoscopic detection of early gastric cancer without blind spots. *Endoscopy* 2019; **51**: 522–31.

15 Hirasawa T, Aoyama K, Tanimoto T *et al*. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer* 2018; **21**: 653–60.

16 Uchiyama K, Takagi T, Naito Y *et al*. Assessment of endoscopic mucosal healing of ulcerative colitis using linked colour imaging, a novel endoscopic enhancement system. *J. Crohns. Colitis* 2017; 1–7.

17 Takeda T, Asaoka D, Nojiri S *et al*. Linked color imaging and the Kyoto Classification of gastritis: evaluation of visibility and inter-rater reliability. *Digestion* 2019. DOI:10.1159/000501534

18 Mori Y, Kudo SE, Philip CW *et al*. Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: an international web-based study. *Endoscopy* 2016; **48**: 1110–8.

19 Misawa M, Kudo SE, Mori Y *et al*. Artificial intelligence-assisted polyp detection for colonoscopy: initial experience. *Gastroenterology* 2018; **154**: 2027–9.

20 Ono S, Kato M, Tsuda M *et al*. Lavender color in linked color imaging enables noninvasive detection of gastric intestinal metaplasia. *Digestion* 2018; **98**: 222–30.