Early detection of oesophageal cancer through colour contrast enhancement for data augmentation

Xiaohong Gao *[a], Stephen Taylor [b], Wei Pang [c] Xin Lu [d], Barbara Braden [e]
[a] Middlesex University, London, UK
[b] University of Oxford, Oxford, UK
[c] Heriot-Watt University, Edinburgh, UK
[d] Ludwig Institute, University of Oxford, Oxford, UK
[e] Translational Gastroenterology Unit, John Radcliff Hospital, University of Oxford, Oxford, UK

ABSTRACT

While white light imaging (WLI) of endoscopy has been set as the gold standard for screening and detecting oesophageal squamous cell cancer (SCC), the early signs of SCC are often missed (1 in 4) due to its subtle change of early onset of SCC. This study firstly enhances colour contrast of each of over 600 WLI images and their accompanying narrow band images (NBI) applying CIE colour appearance model CIECAM02. Then these augmented data together with the original images are employed to train a deep learning based system for classification of low grade dysplasia (LGD), SCC and high grade dysplasia (HGD). As a result, the averaged colour difference (ΔE) measured using CIEL*a*b* increased from 11.60 to 14.46 for WLI and from 17.52 to 32.53 for NBI in appearance between suspected regions and their normal neighbours. When training a deep learning system with added enhanced contrasted WLI images, the sensitivity, specific and accuracy for LGD increases by 10.87%, 4.95% and 6.76% respectively. When training with enhanced both WLI and NBI images, these measures for LGD increases by 14.83%, 4.89% and 7.97% respectively, the biggest increase among three classes of SCC, HGD and LGD. In average, the sensitivity, specificity and accuracy for these three classes are 88.26%, 94.44% and 92.63% respectively for classification of SCC, HGD and LGD, being comparable or exceeding existing published work.

Keywords: Oesophageal cancer, data augmentation, colour contrast, deep learning

1. INTRODUCTION

Oesophagus cancer is the 9th most common cancer [1] and the 6th leading cause of cancer-related death [2]. Globally, the estimated number of new cases was 572,000, of which approximately 509,000 persons died from oesophageal cancer in 2018 [1]. Histologically, there are two major types that constitute the majority of all oesophageal cancers, adenocarcinoma and squamous cell carcinoma cancer (SCC) (87%) [3, 4].

While the overall five-year survival rate of oesophagus cancer is less than 20% [5], this figure can be improved significantly to more than 90% if oesophageal cancer is detected in its intramucosal stage when lymph node metastasis is unlikely and it still can be treated endoscopically. Unfortunately, routine upper gastrointestinal endoscopy carries a significant miss rate for detecting oesophageal cancer and precancerous lesions. Up to 25% [6] of patients diagnosed with oesophageal cancer had a gastroscopy within the previous year are reported as normal findings. Hence there is a clinical urgency to develop a computerised assistant system to improve the detection accuracy of oesophageal pre-cancerous stages, e.g. dysplasia, allowing effective endoscopic treatment and surveillance of affected patients to prevent progression to cancer.

For endoscopy, while light imaging (WLI) has been the gold standard for screening oesophageal squamous cell cancer (SCC). However, the early sign of SCC presents inconspicuous changes in appearance, which results in considerable miss diagnosed rate (1 in 4). Hence, a number of approaches have been proposed to further enhance and thereafter highlight diseases regions. While spectral [7] or multi-spectral imaging (with 4 to 16 mono spectrum bands) or hyper-spectral imaging (16 to 40 bands) techniques have demonstrated potential to depict endogenous contrast capitalising on wavelength-dependent light-tissue interactions [8, 9], these systems require complex designated optical devices to acquire spectral signals, which presents considerable challenges, including physical implementation, prolonged acquisition time and post co-registration.

Another challenge facing the development of a deep learning based decision support system remains the large variations in colour appearance of datasets when obtained from different research centres, leading to that a system usually under performs when testing on an independent cohort of datasets provided from different research centres.

Therefore, this study proposes a novel approach to responds to these challenges. It highlights the early onset of SCC from human perception point of view by increasing the contrast between these concerned regions and their surroundings. Significantly, this contrast enhancement originates from human colour vison by applying a colour appearance model of CIECAM02 [10], standardised by the *Commission Internationale de l'éclairage* (CIE), the first application in endoscopy. CIECAM02 is modelled based on human visual perception to transform between CIE tristimulus values (XYZ) and perceptual attribute correlates by taking into account viewing conditions. These contrasted images are then employed not only to assist clinicians' diagnosis visually but also to be added to training a deep learning based system for detection, delineation and classification of low or high grade dysplasia and cancer as well as in real-time.

2. METHODS

Figure 1 illustrates the steps of colour contrast enhancement for both WLI and narrow band imaging (NBI) images when only two wavelengths of lighting at 415 nm (blue) and 540 nm (green) are employed.

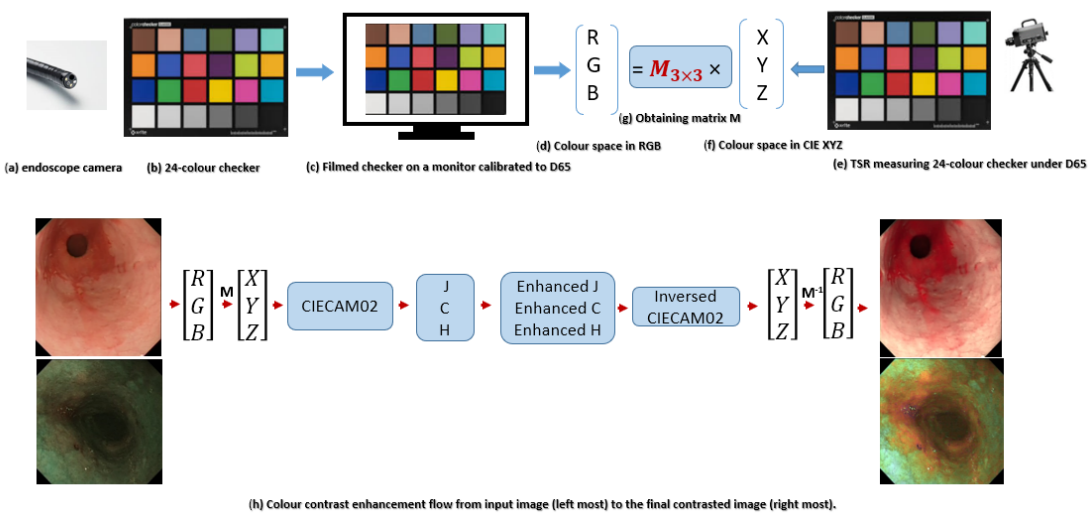

Figure 1. Colour contrast improvement using CIECAM02 model. (a) to (g) Obtaining relationship between image RGB values and CIE tristimulus values XYZ by a 3×3 matrix (M) for endoscopic cameras. (h) Workflow to enhance colour contrast by CIECAM02 model.

The characterisation of endo camera (Figure 1(a)) was firstly performed to establish the relationship by a matrix (M) (Figure 1(g)) between a camera (and monitor) (Figure 1(c)) RGB values (Figure 1(d)) when filming and CIE tristimulus values of XYZ under D65 (average daylight) (Figure 1(f)). XYZ values are a RGB equivalent from human perception's perspective. Then based on CIE XYZ values, colour appearance model CIECAM02 was applied to calculate lightness (J), colourfulness (C) and hue (H), for each pixel of an input image (Figure 1(h), left most). JCH represent the colour attributes from human perception point of view. After enhancing visual meaningful JCH values, they are converted back to XYZ using inversed CIECAM02 model and then to RGB to realise the final contrasted images (Figure 1(h), right most). In Figure 1 (e)(f). the XYZ values for each of 24 colours was measured in advance employing a tele-spectroradiometer (TSR) under D65, whereas the colour monitor to display the video is also calibrated to D65, the average of daylight.

2.1 datasets

Over 600 white light image (WLI) with low grade of dysplasia (LGD) are collected from patents (n=184). All patients included in this study have given written informed consent to donate biopsies. The early sign of LGD (arrows) is confirmed biopsy test, which are also annotated by two experienced endoscopists with at least 15 years of experience.

While the LGD regions have similar hue values with their surroundings, i.e. skin colour within the hue angle range (0-360) between red (H=0) and yellow (H=90), the diseased regions appear to be more colourful, i.e. from skin colour of pinkish to more colourful reddish. Hence, to further increase the differences between suspected regions and their surroundings, the colourfulness is modified to as in Eq. (1).

$$C_{new} = C * C * \beta \tag{1}$$

Where $\beta = max(C)/max(C_{new})$, which allows small colourfulness being smaller and large being larger while ensuring the updated colourfulness in line with the original value range.

After the modification of colour attributes, the JCH values are converted back to XYZ and then to RGB for the final display of enhanced images. Together both original WLI and enhanced ones are employed to train a deep learning system for detection of cancer (SCC), high grade dysplasia (HGD) and low grade dysplasia (LGD).

The overall difference between diseased regions and their immediate surroundings are measured using both CIECAM02 and CIEL*a*b* models and are formulated as Eqs. (2) and (3).

$$\Delta E_{CAM} = \frac{\sum_n \sqrt{(J_d - J_s)^2 + (C_d - C_s)^2 + (H_d - H_s)^2}}{n} \tag{2}$$

$$\Delta E_{Lab} = \frac{\sum_n \sqrt{(L_d - L_s)^2 + (a_d - a_s)^2 + (b_d - b_s)^2)}}{n} \tag{3}$$

Where d refers to diseased region, s the surroundind region and n the total number of diseased-surrounding pairs. The colour attributes, e.g. L*, a*, b*, J, C, H, are the average of the selected each region that is manually delineated. H in Eq. (3) has been normalised to be within [0,100].

2.2 The Architecture of Developed Real Time Processing System

In this study, an end-to-end detection system has been put forward, which built upon the real-time instance segmentation approach of Yolact [11]. The basic underline model was ResNet101 to extract initial feature maps. The program was implemented in Python using the PyTorch libraries. The training takes place by training samples of WLI only (n=666), WLI+ $WLI_{enahnced}$+ NBI (n=2900), and WLI + $WLI_{enhanced}$ + NBI + $NBI_{enhanced}$ (n=4704) for three classes of SCC, HGD, and LGD.

3. RESULTS

Figure 2 illustrates an example of contrast enhancement from the original image (Figure 2(a)) to the colour contrast-enhanced one (Figure 2(b)). The blue and yellow boxes represent the selected suspected and surrounding regions respectively. Figures 2(c) and 2(d) are the detection results for 2(a) and 2(b) respectively where 2(d) presents near correct detection when the developed deep learning system was trained with only LGD class (normal as a default setting if no lesion is detected). The number in the bar on top of the detection box indicates the probability of the classification, e.g. 'suspicious (s:) 0.93' refers to that the detected region has 93% chance to be LGD.

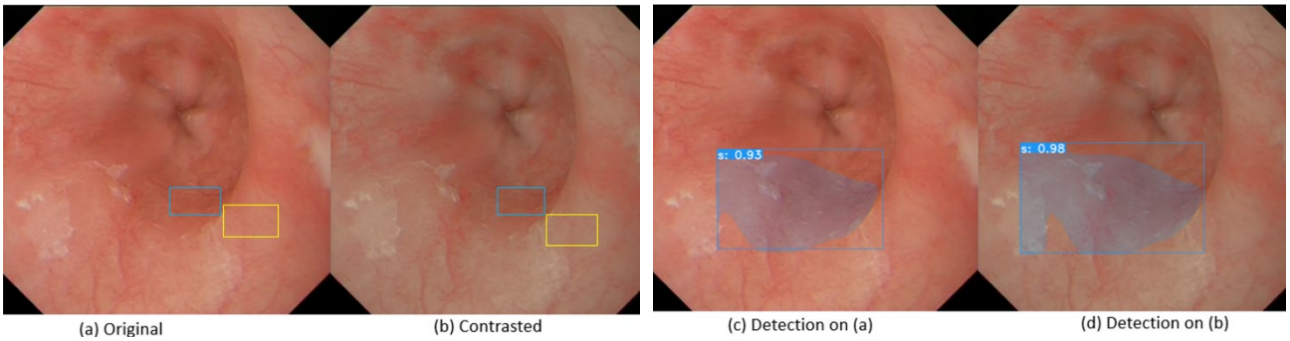

Figure 2. Illustration of colour difference measurement between lesioned region and surroundings (boxes in (a) and (b)) together with LGD detection ((c) and (d), where (d) presents the correct detection for suspected regions when the system is trained with one LGD class (and normal).

Table 1 lists the detection results in terms of sensitivity, specificity, and accuracy trained using WLI with normal data augmentation technique, WLI + contrasted WLI and WLI+BNI+contrasted WLI+contrasted BNI whereas Table 2 measures the colour differences between each LGD region and its surrounding mucosa regions before and after contrast enhancement.

Table 1. The detection results in terms of sensitivity, specificity and accuracy for the deep learning systems trained with WLI, WLI+$WLI_{enhanced}$ and WLI+NBI+$WLI_{enhanced}$ +$NBI_{enhanced}$ images respectively.

Methods	Class	Sensitivity (%)	Specificity (%)	Accuracy (%)
WLI + usual data augmentation	SCC	82.35	95.34	90.50
	HGD	74.19	87.28	84.56
	LGD	74.50	88.29	83.44
	Average	77.13	90.31	86.19
WLI + Enhanced WLI	SCC	86.11	93.08	90.25
	HGD	87.03	90.21	89.49
	LGD	82.97	95.45	90.26
	Average	85.37	93.24	90.47
WLI + NBI + Enhanced WLI + enhanced NBI	SCC	90.47	94.87	93.71
	HGD	85.00	95.12	88.26

	LGD	89.33	93.18	91.41
	Average	**88.26**	**94.44**	**92.63**

Table 2. The colour differences measured using CIEL*a*b* between each LGD and its surrounding mucosa for original WLI and NBI and their enhanced counterparts.

	ΔL	Δa^*	Δb^*	$\Delta E_{L^*a^*b^*}$	ΔL_{CAM}	ΔC_{CAM}	ΔH_{CAM}	$\Delta E_{CAM}(\%)$
WLI	5.76	6.23	7.38	11.60	4.83	9.23	19.30	13.12
Enhanced WLI	7.74	8.48	8.44	14.46	6.62	12.29	27.48	18.35
NBI	8.17	10.59	9.50	17.52	9.57	3.73	6.25	10.82
Enhanced NBI	10.23	16.22	21.56	32.53	12.00	10.06	95.32	33.60

4. NEW BREAKTHROUGH WORK TO BE PRESENTED

This study is probably the first one to enhance colour differences from human perception point of view, which alleviates colour variations between datasets obtained from different research centres because all the enhanced images are processed applying CIECAM02 under a fixed lighting condition, e.g. D65. In addition, it improves detection rate significantly in comparison with usual data augmentation technique by changing colours based on their RGB contents. Significantly, this developed system can operate in real time at 34 frames/second when processing endoscopic videos.

5. CONCLUSION

The challenges facing early detection of oesophageal cancer remain being limited changes presented on the surface of oesophagus and colour variations of datasets obtained from different centres when training an AI-based decision support system. This study takes human colour perception into consideration and pre-process images to enhance image contrast. As a result, not only lesioned regions, mainly suspected or LGD, stand out significantly with colour differences in $\Delta E_{L^*a^*b^*}$ increased from 11.60 to 14.46 and from 13.12 to 18.35 in ΔE_{CAM} for WLI and an increase of 15 and 22 for NBI when measured in $\Delta E_{L^*a^*b^*}$ and ΔE_{CAM} respectively. In addition, with the addition of these enhanced contrasted images to the training, the accuracy improves from 86.19% to 90.25% for WLI and to 92.63% when both WLI and NBI images are applied. These results are based on the independent cohort of test datasets. Significantly, the sensitivity, specificity and accuracy for LGD are increased from 74.50%, 88.29% and 83.44% to 89.33%, 93.18% and 91.42% respectively when both WLI and NBI are included, an improvement by 14%, 5%, and 8% respectively.

In comparison with the work conducted by Osawa et al [7] on colour enhancement based on flexible spectral imaging, e.g.., where the averaged increase of $\Delta E_{L^*a^*b^*}$ is 8.4 units from conventional esophagogastroduodenoscopy (EGD) WLI to NBI, the contrast in our study has been enhanced by 4.74 and 8.9 units for WLI to WLI and WLI to NBI respectively, demonstrating the effectiveness of computational technique on image contrast enhancement.

In comparison with the recently studies on AI-oriented systems [12], this developed system exceeds the state-of-the-art results in relation to early detection of squamous cell neoplasia and is probably the first tangible real-time detection system for endoscopic videos, thanks to the inclusion of contrast enhanced images, by which the sensitivity, specificity and accuracy are 88.2%, 94.4% and 92.6% respectively for classification of three classes, an increase of 11.1%, 4.1% and 6.4% from the outcomes gained without the data enhancement when both WLI and NBI are considered. When only WLI images are employed, the increases are 8.2%, 2.9% and 4.2% respectively.

Another strength of this study is the validation taking place in an independent cohort of patients from another centre, whereas the controls have included real-world patients with Barrett's oesophagus, reflux oesophagitis, candida oesophagitis and anaemia. No magnifying endoscopy is included in this study.

ACKNOWLEDGEMENT

This work was supported by Cancer Research UK [C ref./A 29021] and by the Cancer Research UK (CR-UK) grant number C5255/A18085, through the Cancer Research UK Oxford Centre, and the Oxford Biomedical Research Centre. Their financial support is gratefully acknowledged.

REFERENCES

[1]. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J Clin, American Cancer Society, 2018, 68(6):394–424.

[2]. Pennathur A, Gibson MK, Jobe BA, Luketich JD. Oesophageal carcinoma. Lancet, 2013, 381(9864):400–412.

[3]. Arnold M, Laversanne M, Brown LM, Devesa SS, Bray F. Predicting the Future Burden of Esophageal Cancer by Histological Subtype: International Trends in Incidence up to 2030. Am J Gastroenterol. Nature Publishing Group, 2017, 112(8):1247–55.

[4]. Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. Gut. BMJ Publishing Group, 2015, 64(3):381–7.

[5]. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin. 3rd ed. American Cancer Society, 2014, 64(1):9–29.

[6]. Rodriguez de Santiago E, Hernanz N, Marcos-Prieto HM, et al, Rate of missed oesophageal cancer at routine endoscopy and survival outcomes: a multicentric cohort study**,** United European Gastroenterol J., 2019, 7: 189–198.

[7]. Osawa H, Yamamoto H, Miura Y, Ajibe H, Shinhata H, Yoshizawa M, Sunada K, Toma S, Satoh K, Sugano K. Diagnosis of depressed-type early gastric cancer using small-caliber endoscopy with flexible spectral imaging color enhancement. Dig Endosc. 2012 Jul;24(4):231-6.

[8]. Martinez-Herrera S.E. et al. (2014) Multispectral Endoscopy to Identify Precancerous Lesions in Gastric Mucosa. In: Elmoataz A., Lezoray O., Nouboud F., Mammass D. (eds) Image and Signal Processing. ICISP 2014. Lecture Notes in Computer Science, vol 8509. Springer, Cham, 2014.

[9]. Yoon, J., Joseph, J., Waterhouse, D.J. *et al.* A clinically translatable hyperspectral endoscopy (HySE) system for imaging the gastrointestinal tract, Nat Commun 10,1902, 2019.

[10]. Moroney N, Fairchild MD,. Hunt RWG, Li C, M. Luo MR, Newman T, The CIECAM02 Color Appearance Model. The Tenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications, pp*23-27,* CIC 2002.

[11]. Bolya D., Zhou C., Xiao F., and Lee YJ., YOLACT: Real-time Instance Segmentation, in ICCV, 2019.

[12]. Everson MA. et al., Artificial intelligence for the real-time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: A proof-of-concept study, United Eur. Gastroenterol. J., 2019, 7(2): 297–306.