

Research Paper

Evaluation of Band Selection for Spectrum-Aided Visual Enhancer (SAVE) for Esophageal Cancer Detection

Yen-Chun Chen ^{1,2}, Riya Karmakar ³, Arvind Mukundan ³, Chien-Wei Huang ^{4,5}, Wei-Chun Weng ^{4,✉}, Hsiang-Chen Wang ^{3,6,✉}

1. Department of Gastroenterology, Dalin Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, No. 2, Minsheng Road, Dalin, Chiayi, 62247 Taiwan.
2. School of Medicine, Tzuchi University, No.701, Sec. 3, Zhongyang Rd. Hualien 97004, Taiwan.
3. Department of Mechanical Engineering, National Chung Cheng University, Chia-Yi 62102, Taiwan.
4. Department of Gastroenterology, Kaohsiung Armed Forces General Hospital, 2, Zhongzheng 1st.Rd., Lingya District, Kaohsiung City 80284, Taiwan.
5. Department of Nursing, Tajen University, 20, Weixin Rd., Yanpu Township, Pingtung County 90741, Taiwan.
6. HITSpectra Intelligent Technology Co., Ltd., Kaohsiung City 80661, Taiwan.

✉ Corresponding authors: mikeman7309@gmail.com (W.-C.W.); hcwang@ccu.edu.tw (H.-C.W.).

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <https://ivyspring.com/terms> for full terms and conditions.

Received: 2024.08.24; Accepted: 2024.11.11; Published: 2025.01.01

Abstract

Band selection is a common approach to reduce the data dimensionality of hyperspectral imagery. It extracts several bands of importance in some sense by taking advantage of high spectral correlation. In medical imaging, narrow-band imaging (NBI) is an imaging technique for endoscopic diagnostic medical tests, where light of specific blue and green wavelengths is used to enhance the detail of certain aspects of the surface of the mucosa. A special filter is electronically activated by a switch in the endoscope leading to the use of ambient light of wavelengths of 415 nm (blue) and 540 nm (green). Because the peak light absorption of hemoglobin occurs at these wavelengths, blood vessels will appear very dark, allowing for their improved visibility and in the improved identification of other surface structures. NBI when compared with the white-light imaging (WLI) have proven to have better precision when combined with computer-aided diagnosis (CAD, Intespec C, HITSpectra Intelligent Technology Co., Kaohsiung, Taiwan) in detecting cancerous images. NBI endoscopes are specialized equipment that may not be widely available in all healthcare settings. By leveraging existing WLI endoscopic systems and developing algorithms to simulate NBI imaging, healthcare facilities can achieve similar diagnostic capabilities without the need for additional costly equipment. Therefore, in this study, algorithm known as the SAVE (spectrum-aided visual enhancer) has been developed which can simulate NBI from the WLI images through an intelligent band-selective hyperspectral imaging for Olympus endoscope. The results suggested that the SAVE-NBI images had a better precision and FI-score than the WLI images.

Keywords: esophageal Cancer, hyperspectral imaging, SAVE, dysplasia, SSD, YOLOv5, YOLOv8, narrow-band imaging, white light imaging

Introduction

A hyperspectral imager captures intricate spectral data by utilizing hundreds of narrow bands. [1]. These narrow-band images (NBI) can be in the ultra-violet (UV) [2], visible (VIS), near infra-red (NIR) [3], and mid infra-red (MIR) bands [4]. NBI are currently utilized in various applications such as [5, 6], medical imaging [7-9], remote sensing [10, 11], military [12, 13], agriculture [14, 15] to include a few. Hyperspectral imaging (HSI) is a developing technology that combines traditional imaging and

spectroscopy to capture spatial and spectral data from an object [16]. Hyperspectral images are composed of hypercubes, which are three-dimensional data blocks consisting of two spatial dimensions and one wavelength dimension across the electromagnetic spectrum [17]. HSI generates a substantial amount of data because it has a higher number of bands [18]. Redundancy may exist between bands, with some bands potentially having less discriminative information, leading to increased demands on storage

space, computational load, and communication bandwidth, which are not conducive to real-time applications [19]. It is beneficial to eliminate bands that provide minimal or no distinguishing information.

Band selection is the process of choosing a limited number of hyperspectral bands to eliminate spectral duplication and decrease computational expenses, while still maintaining the important spectral details of objects on the ground [20]. Dimensionality reduction is a crucial preprocessing step for HSI analysis. Its goal is to eliminate spectral redundancy while retaining essential information for future use [21, 22]. Hyperspectral dimensionality reduction can be accomplished by band selection or feature extraction [23]. Many feature extraction methods combine all original bands linearly, which can complicate result interpretation if the chosen criterion is not suitable [24]. However, this study focuses on the band selection. Narrow band imaging (NBI) and Multi-Band Imaging (MBI) are endoscopic imaging techniques that provide real-time visualization of the vascular network and surface texture of the mucosa to aid in tissue characterization, differentiation, and diagnosis [25]. NBI's fundamental principle relies on the absorption of hemoglobin and the wavelength-specific traits of tissue optical properties [26]. Hemoglobin exhibits absorption peaks at 415 nm (blue) and 540 nm (green) in the spectrum, whereas mucosal scattering decreases consistently as wavelength increases [27, 28]. 415 nm light can enhance mucosal capillaries in superficial areas, while 540 nm light improves the visualization of vessels in deeper regions [29, 30].

Previous studies have demonstrated that utilizing NBI in conjunction with CAD yields superior outcomes compared to using white-light imaging (WLI) alone [31-34]. Tatsugami *et al.* assessed the effectiveness of NBI cystoscopy in detecting bladder cancer and examined its diagnostic accuracy in cases of carcinoma *in situ* (CIS) and cases with known urine cytology results [35]. Zhu *et al.*'s study demonstrated that Narrow Band Imaging (NBI) was more effective than White Light Imaging (WLI) in detecting both early and invasive lung cancer [36]. Watanabe *et al.* assessed the effectiveness of laryngoscopy with NBI system for diagnosing precancerous and cancerous laryngeal lesions and concluded that NBI endoscopy is a highly promising diagnostic tool for laryngeal malignant disease [37]. Olympus Medical Systems now offers a commercially available NBI system globally, which serves as an endoscopic diagnostic tool for the gastrointestinal (GI) tract [38]. However, implementing NBI technology necessitates acquiring specialized endoscopes and video processors that are

equipped with NBI capabilities. The high initial equipment cost and continuous maintenance expenses can be substantial, limiting accessibility in certain healthcare environments. NBI technology may not be widely accessible in all healthcare facilities, particularly in resource-constrained settings or smaller clinics that lack standardization [39]. Therefore, an algorithm that is referred to as the SAVE (spectrum-aided visual enhancer, Transfer N, Hitspectra Intelligent Technology Co., Kaohsiung, Taiwan) and a band selection for esophageal cancer detection has been developed in this study. This algorithm has the capability to simulate NBI from the WLI images by utilizing an intelligent band-selective hyperspectral imaging mechanism for the Olympus endoscope.

Material and Methods

Dataset

In this study a dataset consisting of 2761 WLI images, 2761 WLI images were converted to NBI images, and 935 original NBI images used were acquired with the conventional endoscope (CV-290, Olympus) for our investigation. These pictures were resized to 640×640 pixels during preprocessing to avoid possible problems such as limited computer capacity and maintain consistency in format. The process of image annotation was carried out using the widely used LabelImg software, resulting in the creation of an XML file [40]. Although WLI is very effective in diagnosing esophageal cancer, its ability to diagnose dysplasia is limited. Dysplasia often has a visual resemblance to SCC, which makes it difficult to accurately identify with WLI. Thus, the use of HSI conversion methods improves the quality of the images by generating a comprehensive dataset that collects spectrum information across different wavelengths. The overall flow of the study can be seen in Figure 1. It is crucial to emphasize that the annotated dataset will be converted into HSI-NBI pictures utilizing SAVE technology throughout the project. Consequently, three datasets were obtained using WLI, NBI and HSI-NBI images. The dataset was divided into training, testing, and validation sets at a ratio of 70:15:15. Specifically, for the WLI images, 1933 images were used for training, 414 for testing, and 414 for validation. This division ensures a balanced distribution of data across the sets, enabling robust training, accurate testing, and reliable validation of the models used in the study. The WLI images were converted into SAVE images, therefore the same training, validation and testing images was used in the SAVE dataset as well. Moreover, the 935 images of NBI captured was used in the process of creating the

HSI-conversion algorithm and not for the training purpose. In this study, WLI and NBI images were not captured from the same lesion. Unlike other IEE methods, such as FICE, NBI images cannot be taken simultaneously with WLI images. Therefore, WLI images were used to generate HSI-NBI images through the SAVE algorithm, and the actual NBI images served as a reference for comparison. This approach allowed us to evaluate the effectiveness of the SAVE-generated HSI-NBI images in replicating the diagnostic quality of true NBI images. To avoid data leakage, images from the same lesion were kept within a single subset (either training, testing, or validation). This ensured that the model was trained, tested, and validated on independent data, maintaining the integrity of the performance evaluation. The PyTorch deep learning framework was built on the Windows 11 operating system. The Python application was run in Jupyter Notebook, leveraging Python version 3.9.18. The training time of the three models YOLOv5, YOLOv8, and SSD models are 8 hours, 8.5, and 7 hours respectively.

Band selection

To calculate the conversion matrix that converts the RGB image into a hyperspectral image, the spectrometer (model number QE65000 from Ocean Optics) and endoscopic camera must be calibrated. This may be done by giving the spectrometer and endoscope common color targets. This example uses

the x-rite traditional 24-color checker. This color checker was selected because it has six hues of gray and most natural colors. Since endoscope RGB is sRGB, the pictures are converted to CIE 1931 XYZ. The three channels for red, green, and blue will be 0 to 255 before the conversion, then scaled down between 0 and 1. Gamma function was used to transfer color spaces. Mistakes in color separation of filters, color shift, and camera dark current will cause mistakes in this technique. The findings will be wrong due to these errors. To account for these factors, first a matrix expressing the variable V was calculated and performed a regression analysis to get the matrix denoting the correction coefficients C , as shown in equation 1.

$$C = [XYZ_{Spectrum}] \times pinv([V]) \quad (1)$$

The spectrometer provides the values of $XYZ_{Spectrum}$. The XYZ color gamut's Y value is proportional to brightness (0-100) which can be obtained from the light-source spectrum and XYZ color matching function. The $XYZ_{Spectrum}$ begins with the luminance ratio (k), which is calculated from this value after normalization. Equation 2 yields the value of the $XYZ_{Correct}$. It was also found that the $XYZ_{Spectrum}$ and $XYZ_{Correct}$ had 0.5355 RMSE differences making this calibration is necessary. $R_{Spectrum}$ and $XYZ_{Correct}$, the spectrometer reflection and camera corrected spectrum, are assessed. Principle component analysis (PCA) and multiple regression analysis are used to get

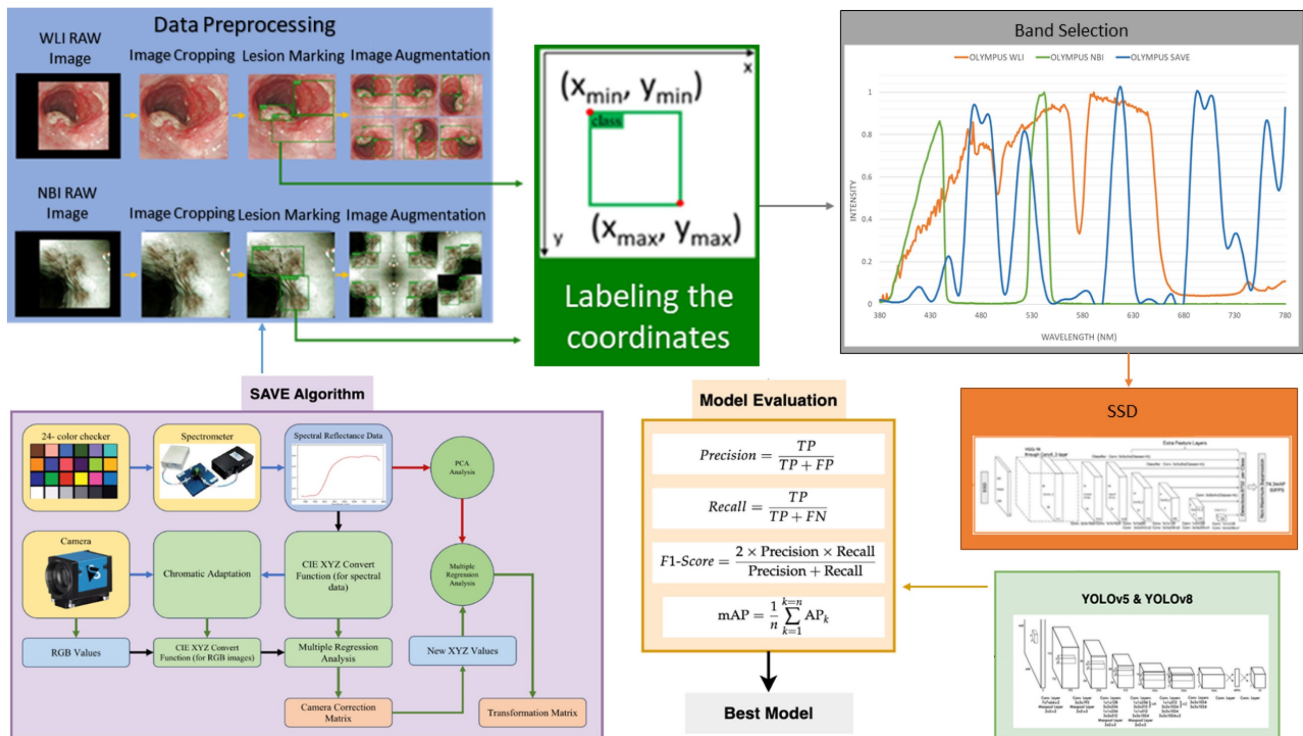


Figure 1. Flowchart of the Process developed in this study.

the conversion matrix (M) and the principal components with the highest significance (PS). V_{Color} was employed in the regression analysis of $XYZ_{Correct}$ and PS since it covers most X, Y, and Z combinations. M may be calculated from equation 3, and equation 4 generates the analog spectrum.

$$(SSpectrum)[XYZCorrect] = [C] x [V] \quad (2)$$

$$M = [Score] x pinv([VColor]) \quad (3)$$

$$SSpectrum = EV[M][VColor] \quad (4)$$

When comparing the final reflection spectrum to the real data, each block's average RMSE was 0.0532, which is insignificant. This was found by comparing the actual and final reflection spectrum values. Thus, using the approach described above, any RGB picture may be converted into a hyperspectral image with a resolution of up to 1 nm and a reflection spectrum identified. The HSI images replicated from the WLI images are named as SAVE (Spectrum-Aided Visual Enhancer) images.

Even though a normal WLI image can be converted to an HSI image, band selection is essential to create a NBI from the HSI images. Luckily, for Olympus endoscope there is a reference NBI capture mode through which an algorithm developed can be compared. Therefore, for Olympus endoscope there is a reference NBI capture mode through which an algorithm developed can be compared. Three factors will affect the color difference between the real NBI and simulated enhanced SAVE images. These factors correspond to the light spectrum, the color-matching function, reflection spectrum. At first, the CIEDE 2000 color difference between the HSI images and the

Olympus endoscope was found. A similar CIEDE 2000 color difference can be found between the HSI images of the capsule endoscopes and other traditional endoscopes with the Olympus NBI image can be found to generate SAVE images for the respective endoscopes making it a versatile application. As shown in Figure 2 the spectrum of the light used in the Olympus WLI endoscope and Olympus NBI images are different. Here, Olympus claims to use 420 nm and 540 nm wavelength filters to achieve the NBI images. However, when NBI images of the Olympus are seen, there seem to be many more bands to make the NBI images more realistic. However, 415 and 540 nm are the peak light absorption of hemoglobin occurs, the lighting spectrum must also be calibrated. It was done using the Cauchy-Lorentz visiting distribution as shown in equation 5.

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]} = \frac{1}{\pi}\left[\frac{\gamma}{(x-x_0)^2+\gamma^2}\right] \quad (5)$$

A 24-color checker is again used for calibrating the real Olympus NBI image with the enhanced SAVE images. For optimizing the lighting spectrum, the dual annealing optimization function is used. This aleatory technique, which was evolved from generalized simulated annealing algorithm which blends the simplified classical simulated annealing (CSA) and fast simulated annealing (FSA), along with an approach for executing a local search based on only certain parameters. Previously the same algorithm has been used for global optimization. The average standard CIEDE 2000 color difference between the 24 color is 5.36 which is negligible. Even though the peak

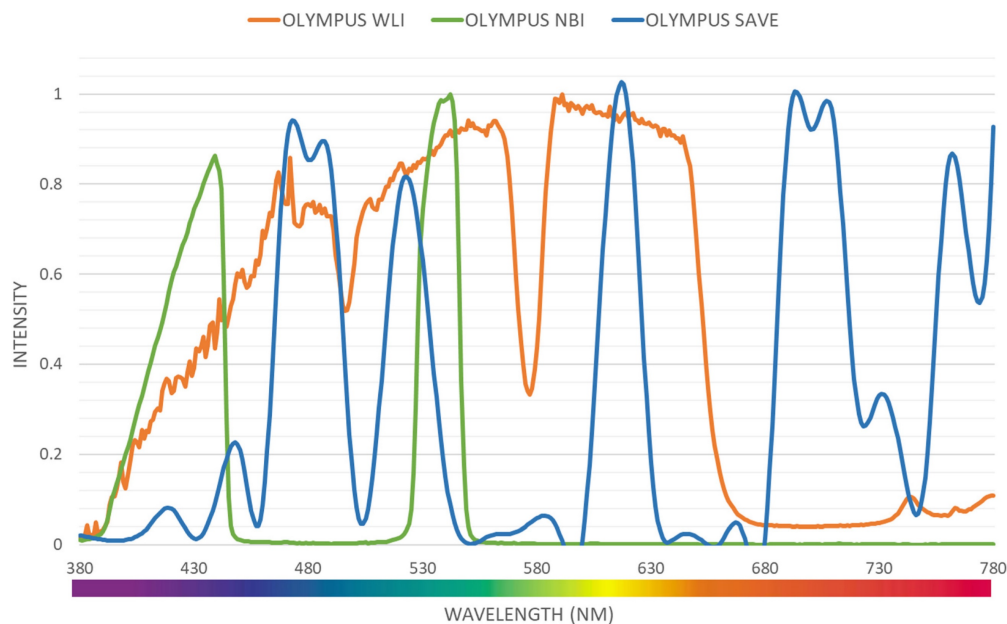


Figure 2. Light intensity of different wavelength for Olympus WLI, Olympus NBI and SAVE images.

absorption wavelength of hemoglobin is 415 and 540 nm, the real NBI image captured by the Olympus endoscope has not only green and blue but shades of brown colors also which corresponds to the wavelength of 650 nm. Therefore, it can be considered that there is a subtle image post-processing which makes the NBI videos more realistic. Hence in this study, apart from 415 and 540 nm three other regions in the wavelength of 480, 600, 700 and 780 nm have also light spectrum as shown in Figure 2.

Machine Learning Model

YOLO

The research used the YOLOv5 D-CNN model developed by Ultralytics [7]. This model employs a grid-based approach, where the input picture is divided into a grid of size $S \times S$. Once an item's precise position is determined inside a grid, the grid assumes the task of detecting the object. The offset is derived by the difference in aspect ratio between the preceding box (Anchor) and the computed value. The grid is classified into either the background or the item category by using a pre-established threshold. Additionally, the technique involves including two adjacent grids located at the middle of the physical frame and aligned with the grid. Ultimately, the combined function of the three grids is crucial in deciding the detection outcome of the present frame. The YOLOv5 model utilizes the displacement of the center position (t_x , t_y) and the proportionate width and height (t_w , t_h) of the forecasted frame in relation to the preceding frame in order to generate predictions. The confidence score and classification probability are calculated simultaneously. The final forecast box probability is determined by performing screening using the IoU threshold and the highest value of NMS. The YOLOv5 neural network architecture consists of four main components: input, feature extraction, feature analysis network, and prediction result output. The input stage employs three techniques to enhance characteristics: mosaic data augmentation, adaptive anchor box computation, and adaptive image scaling. Mosaic data augmentation improves the model's ability to precisely detect and localize small objects inside pictures. Prior to being inputted into the main network, the input data is subjected to a segmentation procedure. The focus structure enhances the channel count of the original three-channel pictures to 12. The CSP architecture reduces parameters and model size by improving the network structure design. The neck module enhances the integration of network features by using the FPN+PAN architecture and integrating bigger feature maps to compensate for the

information degradation at the upper levels of the feature pyramid. The YOLOv5 model utilizes the CIoU_Loss function to evaluate the recognition loss of the identified target rectangle during the prediction process. Ultralytics released YOLOv8 in January 2023 [41]. It is the latest version of the YOLO real-time object recognition models and is much better than the ones that came before it [42]. YOLOv8 improves efficiency by lowering box estimates and speeding up the non-maximum suppression (NMS) process by using an anchor-free design [43]. This is especially helpful when dealing with items that have different forms and aspect ratios. YOLOv8's core is similar to YOLOv5, and it includes the C2f module, which was based on the ELAN idea from YOLOv7. This improves information about gradient flow while keeping the structure light. The SPPF element is kept at the end of the backbone to ensure accuracy on all sizes. YOLOv8 uses the PAN-FPN feature fusion method in the neck section and adds multiple C2f modules as well as a separated head structure that was based on YOLOx. By mixing confidence and regression boxes, they reach a new level of accuracy. Because YOLOv8 is so flexible, it can handle all YOLO versions and switch between them easily, so it can work with a lot of different hardware systems [44]. The network architecture is composed of convolution, batch normalization, and SiLu activation functions.

SSD

The Single-shot multi-box detector (SSD) is a target detector that can recognize many categories in a single shot built using CNN [45, 46]. The SSD used in this research employed a detection framework based on the VGG-16 network. VGG-16 is a deep learning model that has 16 layers, including 13 convolutional layers and 3 fully connected layers. The prediction models' findings were displayed using two methods: displaying the prediction frame and ground truth visually, together with the level of confidence in the forecast. The level of correlation between the predicted outcomes and the annotated location was assessed in relation to the impact of the prediction model on the predicted outcomes. Subsequently, the models' predictive capabilities were assessed using several assessment metrics, including sensitivity, specificity, F1-score, and mean average precision (mAP). The sensitivity of the model reflects its ability to accurately identify symptoms associated with esophageal cancer. The F1-score is a mathematical average that combines precision and recall, providing a reliable measure of the model's performance. mAP, short for mean Average Precision, is a widely used metric for assessing object detection. It quantifies the entire performance of a prediction model by

considering both sensitivity and accuracy.

Results

This work employs a binary classification methodology to identify esophageal cancer by using advanced deep learning algorithms, including YOLOv5, YOLOv8, and SSD. Rather than categorizing images into many classifications indicating various forms of esophageal cancer, the images were grouped into two unique categories: "Cancerous Images" and "Non-Cancerous Images." To enhance implementation and interpretation, we streamlined the categorization job. This technique simplifies the diagnostic procedure, enhancing its accessibility for professionals. The main goal of our study is to distinguish between malignant and benign disorders in the esophagus, which is of great importance in clinical practice. By concentrating on this dichotomous categorization, we provide practical information for healthcare professionals. Acquiring labeled data for many categories of esophageal cancer in this specific scenario is difficult, resulting in a shortage of data or an imbalance in class distribution. By simplifying the problem to a binary classification job, we address these difficulties and improve the model's performance. Binary classification models include intrinsic interpretability in contrast to approaches that involve several classes. The openness is crucial in medical applications, as comprehending the foundation of forecasts is vital for establishing confidence and acceptance. An accurate binary classifier that detects malignant pictures effectively functions as a decision support tool for physicians. This aids in selecting situations for further assessment or action, eventually enhancing patient outcomes. The results of the three models are shown in Table 1.

From Table 1, it can be clearly seen that the SSD model had a much better precision, recall and mAP50 in the HSI spectrum. However, we can clearly, infer from Table 1 that in all the different architectures, the HSI spectrum models performed better than the WLI-RGB models. In this study, the SSD model performed much better than the YOLOv5 and YOLOv8 models. YOLOv5 and YOLOv8 had a precision of 76.4% and 74% in the non-cancerous images in the WLI-RGB model while the HSI spectrum had a precision of 77.1% and 74.8% showing a minimal increase if using the HSI model. However, in the case of cancerous images YOLOv5 model showed an increase of 5.05% while YOLOv8 model showed an increase of 7.45% in precision. While the SSD model showed an increase of 6.4% increase in the non-cancerous images while cancerous images showed an increase of 12.75% increase in precision. Figure 3 shows the comparison between the WLI

image captured by the Olympus endoscope, a similar NBI image also captured by the Olympus endoscope and the SAVE reproduced image based on the WLI images. In the WLI images, the CIE DE76, CIE DE1994, and CIE DE2000 color differences were 65.802, 9.927, and 12.7924, respectively. In comparison, the SAVE images showed significantly larger color differences, with CIE DE76, CIE DE1994, and CIE DE2000 values of 112.36, 54.85, and 40.71, respectively, between normal and cancerous tissue regions. This demonstrates that the SAVE images provide a more pronounced contrast between the lesion and surrounding tissue, making it easier to visually differentiate cancerous regions.

Table 1. Precision, recall, F1-score and mAP50 of YOLOv5, YOLOv8 and SSD models.

Framework	Model	Metrics			
		Precision	Recall	F1-Score	mAP50
YOLOv5	WLI RGB				
	Non-Cancers Images	76.4%	64.9%	70.2%	68.3%
	Cancerous Images	70.7%	55.85%	61.5%	60.9%
	HSI Spectrum				
	Non-Cancers Images	77.1%	63.4%	69.6%	67.5%
	Cancerous Images	75.75%	56.2%	64.35%	60%
YOLOv8	WLI RGB				
	Non-Cancers Images	74%	65.1%	69.3%	68.2%
	Cancerous Images	68%	55%	60%	58.8%
	HSI Spectrum				
	Non-Cancers Images	74.8%	61.4%	67.4%	65.6%
	Cancerous Images	76.45%	51.7%	61.3%	57.6%
SSD	WLI RGB				
	Non-Cancers Images	84.5%	95.533%	88%	75.3%
	Cancerous Images	68%	55%	60%	59%
	HSI Spectrum				
	Non-Cancers Images	90.9%	71.344%	79.9%	78.9%
	Cancerous Images	89.75%	%	89.25%	86.05%

Discussion

SSD surpasses YOLO in this situation due to its use of feature maps at multiple scales, allowing for more effective detection of objects of varying sizes. On the contrary, YOLO operates solely on one level, which may result in limitations. SSD has a briefer training duration than YOLO models. The efficiency of SSD is attributed to its simple design and its one-shot detection technique, which examines the entire image in one go. The YOLO model is efficient in performing inference but requires longer training

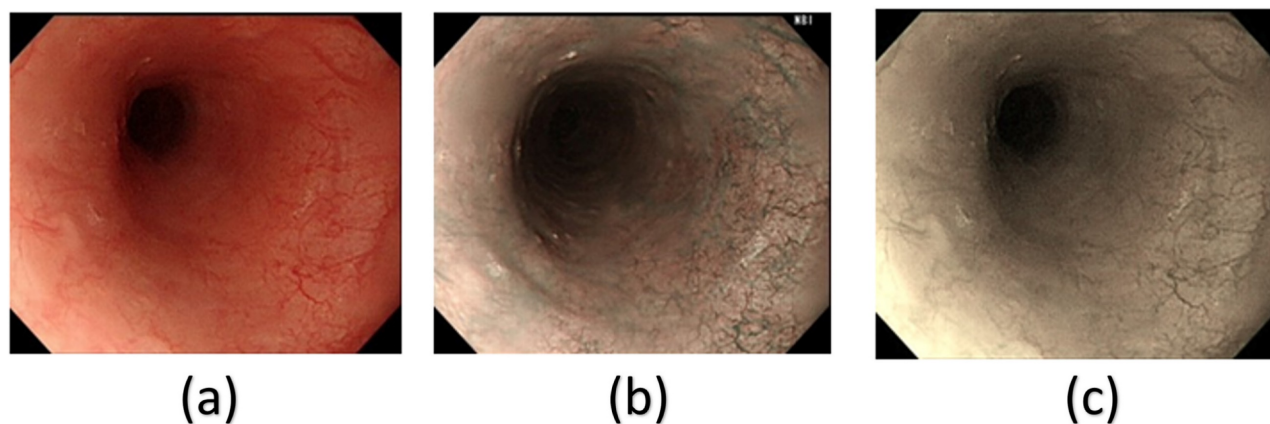


Figure 3. Example of a (a) WLI image from Olympus endoscope, (b) original NBI captured by Olympus endoscope and (c) the enhanced SAVE images obtained from the WLI image.

times due to its complex design. On the contrary, YOLO typically has faster inference rates than SSD because it can analyze the entire image at once, whereas SSD analyzes the image multiple times at different sizes, leading to increased computational complexity. Therefore, the suitability of YOLO designs and SSD model depends on the criteria mentioned above. Moreover, the recall values demonstrate significantly lower values in comparison to the accuracy values. Recall, also known as sensitivity, measures the proportion of correctly identified positive instances out of all actual positive cases by the model. To calculate, divide the number of true positives by the sum of true positives and false negatives. The dataset's significant imbalance, with a large number of negative instances compared to positive ones, can result in an overestimation of accuracy. This occurs because when the model generates a few positive predictions, they are highly probable to be correct, leading to a high accuracy level. The recall was compromised because the model could not accurately identify many positive instances among the majority of negative cases. In the field of medical diagnostics, accuracy should be prioritized over recall. Models were adjusted to reduce the number of false positive predictions, even at the risk of missing some true positive predictions. This compromise led to increased accuracy but decreased comprehensiveness. Regulatory regulations and clinical approval procedures heavily influence the use of models in medical applications. Choosing between YOLO and SSD may also be affected by factors like interpretability, explainability, and validation requirements set by regulatory bodies or clinical standards. Utilizing SSD's multi-scale feature maps and anchor boxes can improve object identification by offering greater clarity, thereby enhancing model interpretability and ensuring regulatory compliance in specific scenarios.

Conclusion

Band selection is a typical hyperspectral images data reduction method. It uses strong spectral correlation to extract numerous important bands. For endoscopic diagnostic medical testing, NBI uses blue and green light to increase mucosa surface detail. A switch in the endoscope activates a filter to utilize ambient light at 415 nm and 540 nm. Because hemoglobin absorbs the most light at these wavelengths, blood vessels look black, improving visibility and surface structure identification. NBI endoscopes are specialized and not ubiquitous in healthcare. Healthcare institutions may attain comparable diagnostic capabilities without expensive equipment by using WLI endoscopic devices and creating algorithms to replicate NBI imaging. This work created the SAVE program to mimic NBI from WLI pictures using intelligent band-selective hyperspectral imaging for Olympus endoscopes. The findings showed that SAVE-NBI photos had superior accuracy and F1-score than WLI images. The SSD model outperformed YOLOv5 and YOLOv8 in this investigation. YOLOv5 and YOLOv8 had 76.4% and 74% accuracy in non-cancerous photos in the WLI-RGB model, whereas the HSI spectrum had 77.1% and 74.8%, a modest gain. YOLOv5 and v8 models increased accuracy by 5.05% and 7.45%, respectively, for malignant pictures. The SSD model increased accuracy by 6.4% in non-cancerous pictures and 12.75% in malignant images.

Supplementary Material

Supplementary methods, figures and information.
<https://www.jcancer.org/v16p0470s1.pdf>

Acknowledgements

Funding

This research was supported by the National Science and Technology Council, The Republic of China under the grants NSTC 113-2221-E-194-011-MY3. This work was financially/partially supported by the Dalin Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation-National Chung Cheng University Joint Research Program DTCRD113-C-02, and Kaohsiung Armed Forces General Hospital research project KAFGH_D_113041 in Taiwan.

Author contributions

All authors have read and agreed to the published version of the manuscript. Conceptualization, Y.-C.C., A.M., T.-H.C., W.-C.W. and H.-C.W., methodology, Y.-C.C., A.M., and R.K.; software, A.M. and R.K.; validation, Y.-C.C., A.M., T.-H.C., W.-C.W. and H.-C.W.; formal analysis, R.K., and A.M.; investigation, R.K., A.M.; resources, H.-C.W., and R.K. ; data curation, A.M., and C.-W.H.; writing—Original draft preparation, A.M.; writing—Review and editing, A.M, T.-H.C., W.-C.W. and C.-W.H.; supervision, A.M, Y.-C.C., H.-C.W. ; project administration, Y.-C.C. and H.-C.W.

Data availability statement

The data presented in this study are available in this article upon considerable request to the corresponding author (H.-C. W.).

Informed consent statement

Written informed consent was waived in this study because of the retrospective, anonymized nature of study design.

Institutional review board statement

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of the Institutional Review Board of Kaohsiung Armed Forces General Hospital (KAFGHIRB 112-018) and the Institutional Review Board of Dalin Tzu Chi General Hospital (B11301023).

Competing Interests

The authors have declared that no competing interest exists.

References

- Lu B, Dao PD, Liu J, He Y, Shang J. Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture. *Remote Sensing*. 2020; 12: 2659.
- Al Ktash M, Stefanakis M, Boldrini B, Ostertag E, Brecht M. Characterization of Pharmaceutical Tablets Using UV Hyperspectral Imaging as a Rapid In-Line Analysis Tool. *Sensors*. 2021; 21: 4436.
- Mukundan A, Tsao YM, Wang HC. Design, Development, and Validation of a Portable and Low-Cost Visible and Near-Infrared Hyperspectral Imaging Module. In *Proceedings of the International Conference on Mechanical, Automotive and Mechatronics Engineering (ICMAME 2023)*, Dubai, UAE; 2023.
- Kniazev K, Zaitsev E, Zhang S, Ding Y, Ngo L, Zhang Z, et al. Hyperspectral and Nanosecond Temporal Resolution Widefield Infrared Photothermal Heterodyne Imaging. *ACS Photonics*. 2023; 10: 2854-60.
- Mukundan A, Tsao Y-M, Cheng W-M, Lin F-C, Wang H-C. Automatic Counterfeit Currency Detection Using a Novel Snapshot Hyperspectral Imaging Algorithm. *Sensors*. 2023; 23: 2026.
- Mukundan A, Tsao Y-M, Wang H-C. Detection of counterfeit holograms using hyperspectral imaging. *Holography, Diffractive Optics, and Applications XIII: SPIE*; 2023. 17-25.
- Huang H-Y, Hsiao Y-P, Mukundan A, Tsao Y-M, Chang W-Y, Wang H-C. Classification of Skin Cancer Using Novel Hyperspectral Imaging Engineering via YOLOv5. *Journal of Clinical Medicine*. 2023; 12: 1134.
- Wang C-Y, Mukundan A, Liu Y-S, Tsao Y-M, Lin F-C, Fan W-S, et al. Optical Identification of Diabetic Retinopathy Using Hyperspectral Imaging. *Journal of Personalized Medicine*. 2023; 13: 939.
- Yang K-Y, Mukundan A, Tsao Y-M, Shi X-H, Huang C-W, Wang H-C. Assessment of hyperspectral imaging and CycleGAN-simulated narrowband techniques to detect early esophageal cancer. *Scientific Reports*. 2023; 13: 20502.
- Mukundan A, Huang C-C, Men T-C, Lin F-C, Wang H-C. Air Pollution Detection Using a Novel Snap-Shot Hyperspectral Imaging Technique. *Sensors*. 2022; 22: 6231.
- Mukundan A, Hong-Thai N, Wang H-C. Detection of PM 2.5 Particulates using a Snap-Shot Hyperspectral Imaging Technology. *Proceedings of the 2022 Conference on Lasers and Electro-Optics Pacific Rim*. Sapporo: Optica Publishing Group; 2022.
- Shimoni M, Haelterman R, Perneel C. Hyperspectral imaging for military and security applications: Combining myriad processing and sensing techniques. *IEEE Geoscience and Remote Sensing Magazine*. 2019; 7: 101-17.
- Briottet X, Boucher Y, Dimmeler A, Malaplate A, Cini A, Diani M, et al. Military applications of hyperspectral imagery. *Targets and backgrounds XII: Characterization and representation: SPIE*; 2006. 82-9.
- Adão T, Hruška J, Pádua L, Bessa J, Peres E, Morais R, et al. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote sensing*. 2017; 9: 1110.
- Dale LM, Thewis A, Boudry C, Rotar I, Dardenne P, Baeten V, et al. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review. *Applied Spectroscopy Reviews*. 2013; 48: 142-59.
- Gowen AA, O'Donnell CP, Cullen PJ, Downey G, Frias JM. Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*. 2007; 18: 590-8.
- Lu R, Chen Y-R. Hyperspectral imaging for safety inspection of food and agricultural products. *Pathogen Detection and Remediation for Safe Eating: SPIE*; 1999.121-33.
- Burger J, Gowen A. Data handling in hyperspectral image analysis. *Chemometrics and Intelligent Laboratory Systems*. 2011; 108: 13-22.
- Guo B, Gunn SR, Damper RL, Nelson JDB. Band Selection for Hyperspectral Image Classification Using Mutual Information. *IEEE Geoscience and Remote Sensing Letters*. 2006; 3: 522-6.
- Sun W, Du Q. Hyperspectral band selection: A review. *IEEE Geoscience and Remote Sensing Magazine*. 2019; 7: 118-39.
- Harsanyi JC, Chang C-I. Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. *IEEE Transactions on geoscience and remote sensing*. 1994; 32: 779-85.
- Lunga D, Prasad S, Crawford MM, Ersoy O. Manifold-learning-based feature extraction for classification of hyperspectral data: A review of advances in manifold learning. *IEEE Signal Processing Magazine*. 2013; 31: 55-66.
- Miao X, Gong P, Swope S, Pu R, Carruthers R, Anderson GL. Detection of yellow starthistle through band selection and feature extraction from hyperspectral imagery. *Photogrammetric Engineering and Remote Sensing*. 2007; 73 (9): 1005 - 1015.
- Martínez-Usó Martínez-Usó A, Pla F, Sotoca JM, García-Sevilla P. Clustering-based hyperspectral band selection using information measures. *IEEE Transactions on Geoscience and Remote Sensing*. 2007; 45: 4158-71.
- Song LMWK, Adler DG, Conway JD, Diehl DL, Farraye FA, Kantsevov SV, et al. Narrow band imaging and multiband imaging. *Gastrointestinal endoscopy*. 2008; 67: 581-9.
- Du Le VN, Wang Q, Gould T, Ramella-Roman JC, Joshua Pifer T. Vascular contrast in narrow-band and white light imaging. *Appl Opt*. 2014; 53: 4061-71.
- Curvers W, Bohmer C, Mallant-Hent R, Naber A, Ponsioen C, Ragunath K, et al. Mucosal morphology in Barrett's esophagus: interobserver agreement and role of narrow band imaging. *Endoscopy*. 2008; 40: 799-805.
- Takano JH, Yakushiji T, Kamiyama I, Nomura T, Katakura A, Takano N, et al. Detecting early oral cancer: narrowband imaging system observation of the oral mucosa microvasculature. *International journal of oral and maxillofacial surgery*. 2010; 39: 208-13.
- Gono K. An Introduction to High-Resolution Endoscopy and Narrowband Imaging. *Comprehensive Atlas of High-Resolution Endoscopy and Narrowband Imaging*. 2017: 7-15.

30. Wang Q, Le D, Ramella-Roman J, Pfefer J. Visualization of mucosal vasculature with narrow band imaging: a theoretical study. *Design and Quality for Biomedical Technologies V: SPIE*; 2012. 18-26.
31. Kurumi H, Nonaka K, Ikebuchi Y, Yoshida A, Kawaguchi K, Yashima K, et al. Fundamentals, Diagnostic Capabilities, and Perspective of Narrow Band Imaging for Early Gastric Cancer. *Journal of clinical medicine*. 2021; 10: 2918.
32. Kraft M, Fostiropoulos K, Gürtler N, Arnoux A, Davaris N, Arens C. Value of narrow band imaging in the early diagnosis of laryngeal cancer. *Head & neck*. 2016; 38: 15-20.
33. Herth FJ, Eberhardt R, Anantham D, Gompelmann D, Zakaria MW, Ernst A. Narrow-band imaging bronchoscopy increases the specificity of bronchoscopic early lung cancer detection. *Journal of Thoracic Oncology*. 2009; 4: 1060-5.
34. Piazza C, Cocco D, De Benedetto L, Del Bon F, Nicolai P, Peretti G. Narrow band imaging and high definition television in the assessment of laryngeal cancer: a prospective study on 279 patients. *European Archives of Oto-rhino-laryngology*. 2010; 267: 409-14.
35. Tatsugami K, Kuroiwa K, Kamoto T, Nishiyama H, Watanabe J, Ishikawa S, et al. Evaluation of narrow-band imaging as a complementary method for the detection of bladder cancer. *Journal of endourology*. 2010; 24: 1807-11.
36. Zhu J, Li W, Zhou J, Chen Y, Zhao C, Zhang T, et al. The diagnostic value of narrow-band imaging for early and invasive lung cancer: a meta-analysis. *Clinics*. 2017; 72: 438-48.
37. Watanabe A, Taniguchi M, Tsujie H, Hosokawa M, Fujita M, Sasaki S. The value of narrow band imaging for early detection of laryngeal cancer. *European archives of oto-rhino-laryngology*. 2009; 266: 1017-23.
38. Muto M, Horimatsu T, Ezoe Y, Hori K, Yukawa Y, Morita S, et al. Narrow-band imaging of the gastrointestinal tract. *Journal of gastroenterology*. 2009; 44: 13-25.
39. Emura F, Saito Y, Ikematsu H. Narrow-band imaging optical chromocolonoscopy: advantages and limitations. *World journal of gastroenterology: WJG*. 2008; 14: 4867.
40. Yakovlev A, Lisovychenko O. An approach for image annotation automatization for artificial intelligence models learning. *Адаптивні системи автоматичного управління*. 2020; 1: 32-40.
41. Wen G, Li M, Luo Y, Shi C, Tan Y. The improved YOLOv8 algorithm based on EMSPConv and SPE-head modules. *Multimedia Tools and Applications*. 2024: 1-17.
42. A. Mehra. (2023). Understanding YOLOv8 Architecture, Applications and Features. Accessed: November 9, 2024. [Online]. Available:<https://www.labellerr.com/blog/understanding-yolov8-architecture-reapplications-features>.
43. Sohan M, Sai Ram T, Reddy R, Venkata C. A Review on YOLOv8 and Its Advancements. *International Conference on Data Intelligence and Cognitive Informatics*: Springer; 2024. 529-45.
44. Chabi Adjoko E, Sanda Mahama AT, Gouton P, Tossa J. Automatic localization of five relevant Dermoscopic structures based on YOLOv8 for diagnosis improvement. *Journal of Imaging*. 2023; 9: 148.
45. Kumar A, Zhang ZJ, Lyu H. Object detection in real time based on improved single shot multi-box detector algorithm. *EURASIP Journal on Wireless Communications and Networking*. 2020; 2020: 1-18.
46. Ning C, Zhou H, Song Y, Tang J. Inception single shot multibox detector for object detection. *2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*: IEEE; 2017. 549-54.