

Article

An Improved Method of Polyp Detection Using Custom YOLOv4-Tiny

Mukhtorov Doniyorjon ^{1,2}, Rakhmonova Madinakhon ¹, Muksimova Shakhnoza ¹  and Young-Im Cho ^{1,2,*} 

¹ Department of IT Convergence Engineering, Gachon University, Sujeong-gu, Seongnam-si 461-701, Korea

² Department of Computer Engineering, Gachon University, Sujeong-gu, Seongnam-si 461-701, Korea

* Correspondence: yicho@gachon.ac.kr

Abstract: Automatic detection of Wireless Endoscopic Images can avoid dangerous possible diseases such as cancers. Therefore, a number of articles have been published on different methods to enhance the speed of detection and accuracy. We also present a custom version of the YOLOv4-tiny for Wireless Endoscopic Image detection and localization that uses a You Only Look Once (YOLO) version to enhance the model accuracy. We modified the YOLOv4-tiny model by replacing the CSPDarknet-53-tiny backbone structure with the Inception-ResNet-A block to enhance the accuracy of the original YOLOv4-tiny. In addition, we implemented a new custom data augmentation method to enhance the data quality, even for small datasets. We focused on maintaining the color of medical images because the sensitivity of medical images can affect the efficiency of the model. Experimental results showed that our proposed method obtains 99.4% training accuracy; compared with the previous models, this is more than a 1.2% increase. An original model used for both detection and the segmentation of medical images may cause a high error rate. In contrast, our proposed model could eliminate the error rate of the detection and localization of disease areas from wireless endoscopic images.

Keywords: wireless capsule endoscopy (WCE); polyp; classification; neural networks; localization



Citation: Doniyorjon, M.; Madinakhon, R.; Shakhnoza, M.; Cho, Y.-I. An Improved Method of Polyp Detection Using Custom YOLOv4-Tiny. *Appl. Sci.* **2022**, *12*, 10856. <https://doi.org/10.3390/app122110856>

Academic Editor: Christian W. Dawson

Received: 13 September 2022

Accepted: 22 October 2022

Published: 26 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The importance of the early detection and diagnosis of dangerous diseases has risen sharply in medical healthcare and, simultaneously, AI-based diagnostic systems have had a tremendous positive impact on the real-time detection and localization of disease areas. Among several reliable methods, Computer Vision (CV) techniques have also become common in real-time detection. Gastrointestinal cancer is very dangerous and common among patients, and early detection of the cancer using a high-quality, reliable AI-based model is crucial. Learning of Wireless Endoscopic Images allows detection of even small diseases to avoid dangerous effects. Most studies are related to computer vision-based diagnostic tools for the detection of tumors, polyps [1], malignancy [2], and ulcers [3]. Polyps are particularly difficult to detect [1,4], since they are smaller than 10 mm and have been shown to be most missed during diagnosis [5]. There is a high demand for endoscopic examination and it requires both equipment and qualified experts for diagnostic analysis for early detection, which may result in a better understanding of cancer screening and reduce deaths due to cancer. Over the past few years, new methods for reducing human error have been developed [6].

Some of the more recent research has already achieved high-quality results for both the real-time endoscopic detection and classification of cancer using several techniques, such as machine learning (ML) [7] techniques, used in wireless capsule endoscopy (WCE). Furthermore, pixel-based models, such as bleeding-detection models, have also been effectively used in the GI tract [8]. Deep learning is becoming widely recognized as a powerful technique and several deep learning methods [9,10] have been developed, such as AlexNet [11] and VGG16 [12] for detecting bleeding regions, GoogleNet [13] for polyp classification,

and Inceptionv4 [14] for polyp localization. Several image-segmentation models have also been developed for polyp segmentation [15] and pancreas segmentation [16]. The U-Net network, with a typical encoder and decoder, was the main model used in medical image segmentation. UNet++ [17] was developed by adding an extra layer to reduce the number of features, which resulted in improved disease segmentation. With the increasing significance of polyp classification, ResUNet++ [18] was developed as a U-Net-shaped network with a residual block.

Recently, hybrid models have emerged as a promising method for the detection and classification of tumors. These models include not only the localization of abnormal areas, but also the detection of bleeding, lesions, and polyps. Techniques based on computer vision and deep learning, such as Grad-Cam++ [19] and custom SegNet [20], have achieved good results by using a convolutional neural network (CNN) classifier before the segmentation of the disease area.

In this study, we developed a new custom object detection model, which is a modified version of the YOLOv4-tiny model, where the Inception-ResNet-A block replaced the backbone architecture of the YOLOv4-tiny, is replaced with a new block network for high-accuracy disease classification and disease detection. In addition, since we used the small open-source KVASIR [21] dataset, we applied efficient data augmentation operations to enhance the dataset.

2. Related Work

Existing research on medical image analysis, particularly wireless endoscopic image analysis, is widely used for the detection of bleeding [10], polyp [1,22], malignancy [2], ulcers [23], and gastrointestinal tumors [24]. Color-based augmentation plays a crucial role in converting RGB images [10]. Shape and texture analyses show different performances for different disease classes, such as polyps, ulcers, and tumors. Wavelet transform [25], and a bag of visual words [26] are examples of new approaches that use various types of data features. Several threshold-based image classification techniques using ML algorithms have been developed [13,14,27]. Deep sparse SVM (DSSVM) [28] has been developed for identifying abnormal frames. The DSSVM is trained on color and texture features extracted from image super-pixels. Recently, deep learning has greatly improved medical diagnosis, with the most commonly used CNN models, such as AlexNet [11], Google Net [13], and VGG16 [12], used for wireless endoscopic image analysis for the early detection of endoscopic diseases. Deep learning techniques for automatic detection of intestinal bleeding from capsule endoscopy images [10] used the SegNet [13] architecture for the segmentation of the disease area in endoscopic images, particularly the bleeding area. The model could detect only a single class, bleeding, from endoscopic images; it scored above 90% accuracy in segmenting areas of bleeding in endoscopic images. Another related study developed a method for the automatic detection of bowel diseases in endoscopy images [7]. This method was divided into two main categories, detection and classification, for diagnosing gastrointestinal bleeding, polyps, and tumors. Feature extraction techniques, such as the Laplacian of Gaussian, were used for detection, and multilayer perceptron based on back-propagation was used for classification. This method performed well with a multilayer perceptron neural network with no prior knowledge of the statistical distribution. For practical implementation, the goal should have been to collect a more accurate dataset to obtain a high-performance score. A CNN model [22] was developed for polyp recognition from endoscopy images. This model used a stacked auto-encoder with an image manifold constraint to recognize polyps. Furthermore, they tried asymmetrical and unsupervised learning techniques for learning image features automatically; however, the model had low performance because of overfitting. A similar work [27] developed an enhanced CNN-based gaze estimation model for wireless capsule endoscopy images, in which an auto-encoder CNN model predicted visual saliency map estimation; this reconstructed model efficiently detected Gaza patterns.

Diagnosis from medical images has always been a difficult task because it requires detection, localization, classification, and other related tasks. While using wireless endoscopic images, a hybrid model was used to obtain accurate results in the early detection of polyps that could cause cancer, and in the classification of gastrointestinal diseases. A CNN model for the real-time detection of early esophageal neoplasia [24] was developed using Inception-ResNetv2 for feature extraction and YOLOv2 [29] for lesion detection. Between 2016 and 2018, electronic endoscopy was used to collect 916 images from 30 patients. The model identified 24 of the 26 patients as having dysplasia with 92.3% accuracy; however, it was not tested on other real endoscopic images and showed low performance in real-time multi-detection of the disease. ResUNet++ was applied successfully in polyp segmentation [30], and conditional random field (a class of statistical models) and test-time augmentation [18]. Deep learning architecture has achieved high performance in wireless endoscopic image analysis. The main goal of these models is related to the segmentation of the polyp area. They used cross-date testing techniques to determine the capacity of the model. Furthermore, these models easily handle small, medium, and large polyps. Performance accuracy is greater than 93%. Several limitations are associated with the segmentation of polyps, such as bowel qualification and camera angle. Recently, deep networks have also been used for medical image detection and localization in endoscopy images [31], which is a hybrid model that solves both the detection and localization of abnormal images while also providing the disease-segmented area. They used two custom models: SegNet [13] and Grad-CAM++ [19]. The model dataset consisted of 1025 wireless endoscopic images of KID [28] and the KVASIR dataset [21]. As the model is quite similar to our proposed method, except for the model structure and the CNN model architecture, our model provides classification and localization rather than segmentation of fixtures.

3. Dataset

The initial version of the KVASIR dataset [21] consists of a total of 4000 images with eight classes, of which the first three classes are normal images (Figure 1) and the other five classes are abnormal images (Figure 2). The second version of the dataset includes the total number of 8000 images that differed between 720×576 and 1920×1072 image resolution.

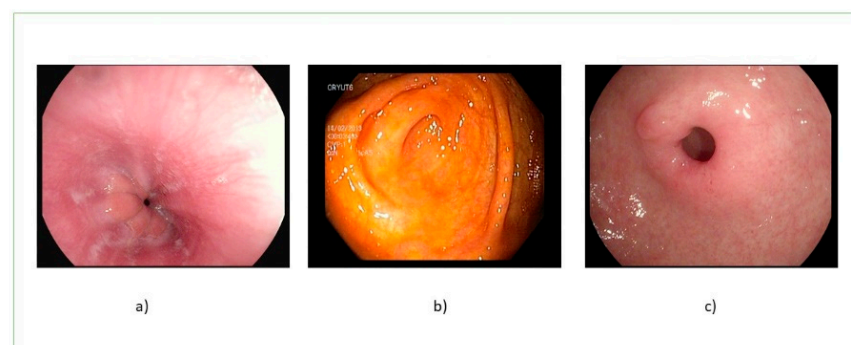


Figure 1. Normal images: (a) normal z-line, (b) normal cecum, and (c) normal pylorus.

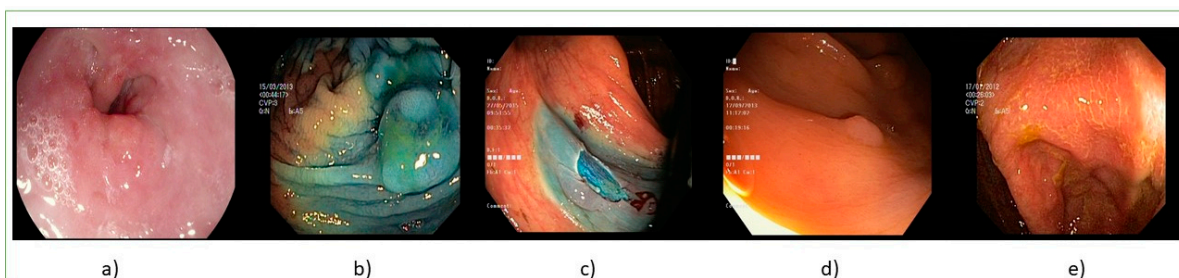


Figure 2. Abnormal images: (a) esophagitis, (b) dyed, and lifted polyps, (c) dyed dissection margins, (d) polyps, and (e) ulcerative colitis.

4. Proposed Architecture

4.1. Overall Architecture of the Proposed Model

In the initial stage, we planned to use only the custom YOLO version to detect polyps, but we did not expect to encounter some missing data while doing so because the gastrointestinal tract is home to various diseases. To address these issues, we decided to develop a new approach that enhances the model architecture and reduces the number of missing polyps during detection. YOLO, proposed by Redmon et al. [29], has been widely used for real-time object detection. Compared with the other models, the YOLO series model simplifies the object detection task into a single regression task and can therefore detect objects in real time. YOLOv4 [32] is an improved version of YOLOv3 [33]. Furthermore, YOLOv4-tiny [5] is a modified version of YOLOv4 that is simpler and faster than the original version. The number of features that could be used for classification and regression detection was reduced from 60 million to 6 million with the application of the CSPOSANet (CSP) backbone.

In this study, we propose a new custom model for wireless endoscopic image analysis based on the improved YOLOv4-tiny model with the Inception-ResNet-A block. We replaced the previous CSP module [34] with our custom CNN block to increase the classification accuracy for the detection and localization of polyp areas. Figure 3 shows the overall architecture of our proposed model architecture. The original version of the YOLOv4 backbone is the Darknet53-tiny, which receives a $608 \times 608 \times 3$ input size and sends it to the $32 \times 3 \times 3$ convolutional layer in the first step. In the next step, three CSP blocks come with max pooling of the baseline SCP block, which is shown in the next part of Figure 3. The CSP block divides the features of the base layer into two parts; the main function of the first part is to directly form the residual edge, and the final output can be obtained by convolution using the splicing of the second part. This network structure can reduce computational complexity while maintaining network performance.

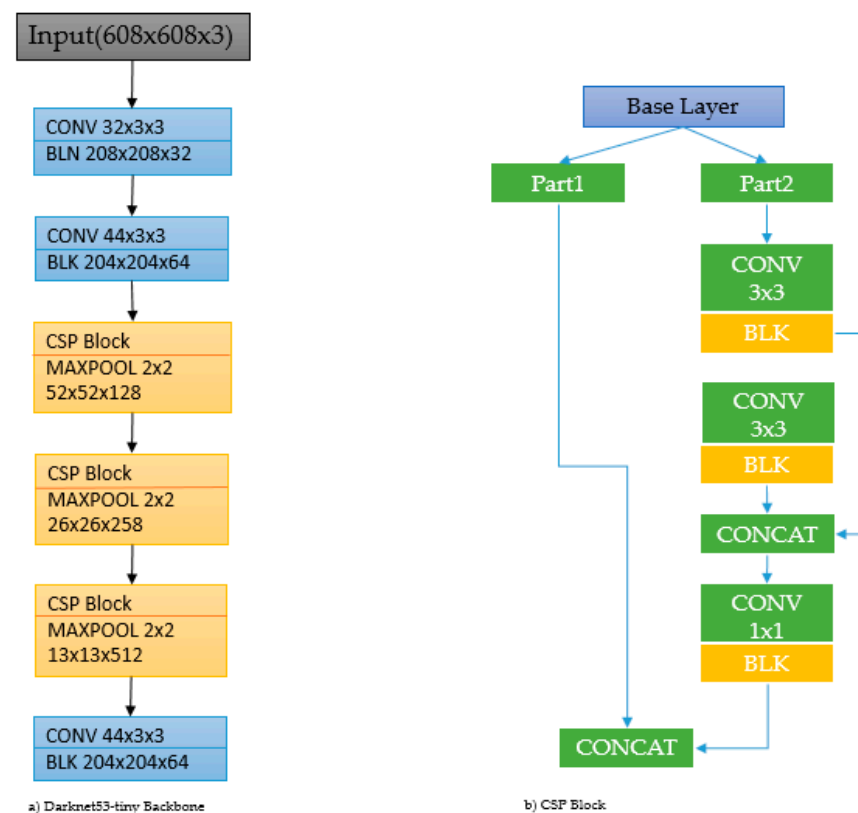


Figure 3. Backbone of original YOLOv4-tiny. Left side: Darknet53-tiny backbone architecture. Right side: CSP block.

Inception-ResNet-A block was modified by combining one of the blocks of the traditional Inception-ResNet architecture [35] with the ResNet model [36]. The ResNet and Inception blocks can be used in combination because the ResNet module consists of residual connections that can be used to train a deep neural network, while the inception block can handle more information from input images. Generally, the architectures of several versions of Inception-ResNet, such as A-, B-, and C-modularized Inception-ResNet blocks, the original CSPOSA block, and the updated Inception-ResNet-A block, include stem blocks as shown in Figure 4. Inception-ResNet-A block is modified as Inception 1×1 convolution block connected with ReLU activation.

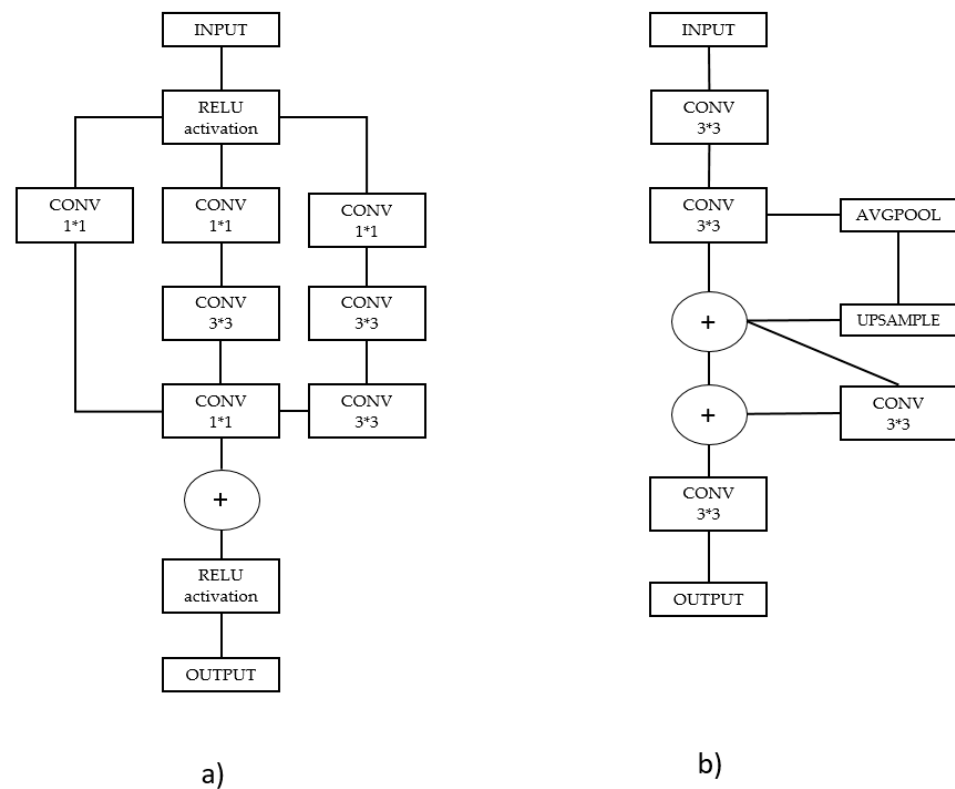


Figure 4. (a) Inception-ResNet-A block and (b) original CSP Block.

4.2. A New Approach to Data Augmentation

We used a custom data augmentation method for increasing the dataset size and performed operations, such as normalization, adjusting brightness, and hue augmentation. Since medical images are sensitive, we only made small changes to the images. Normalization was performed for the training dataset to scale data points for shifting and rescaling. In addition, standardization is another scaling method used where the values are around the mean value with a unit standard deviation. Table 1 with Formula (1) represents the *STD* value scaled between 0 and 1. μ and σ are the means of the given distribution and standard deviation of the given distribution, respectively:

$$STD = \frac{x - \mu}{\sigma} \quad (1)$$

In Equation (3), $\alpha > 0$ and β is a parameter that affects the brightness and contrast. $F(x)$ is the source of pixel intensity and $g(x)$ is the output of pixel intensity, respectively.

Table 1. Standard deviation and mean for the training dataset.

STD	Mean
0.4857, 0.3460, 0.2983	0.2998, 0.2280, 0.2042

We decided not to use the hue automatically, because if the color of the image changes dramatically, overfitting will occur during training. However, brightness was chosen at random with a hue > 2. In Equation (2), $\alpha > 0$ and β is a parameter that affects the brightness and contrast. $F(x)$ is the source of pixel intensity and $g(x)$ is the output of pixel intensity, respectively:

$$g(x) = \alpha(f(x) + \beta) \quad (2)$$

Secondly, we rotated all food images to 90° , 180° , and 270° and randomly rotated the food images to an angle value greater than 20° and chosen randomly:

$$\text{new}(X) = \cos(\text{angle})x - \sin(\text{angle})y \quad (3)$$

$$\text{new}(Y) = \sin(\text{angle})x + \cos(\text{angle})y \quad (4)$$

The above Equations (3) and (4) represent a mathematical form of random rotation. New X and Y coordinates are obtained from the formula with a random angle. Figure 5 shows the images before and after data augmentation.

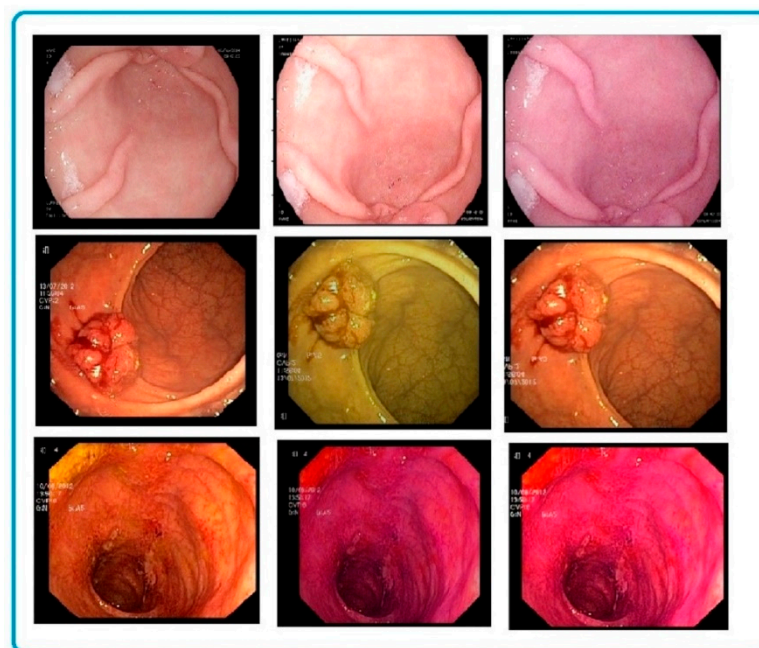


Figure 5. Data augmentation visualization for custom CNN models using random rotation, hue, and brightness. The left side consists of the original dataset, and data augmentation results from the second part of the table.

5. Experimental Results

To train and test the models to compare with each other, we used Visual Studio 2019 C++ on a PC with an AMD Ryzen 12-core processor of 3.70 GHz, 64 GB RAM, and an Nvidia GeForce RTX 3090 graphics card. We compared the accuracy of our model with that of the previous methods using the metrics of precision, recall, and F-score. Precision indicates the number of positive predictions that are correct as shown in Equation (5),

whereas recall (Equation (6)) indicates the number of positive cases correctly predicted by the classifier:

$$\text{precision} = \frac{TP}{TP + FP} \quad (5)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (6)$$

where TP, FN, and FP denote the true positive, false negative, and false positive, respectively. F1 score is calculated based on precision and recall as shown in Equation (7):

$$F1 = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \quad (7)$$

A high F1 score indicates a better performance [37]. The F1-score metric was used to compare the performance of our model with the previous models as presented in Figure 6. It is observed that our model performs better than the other models during both training and testing.

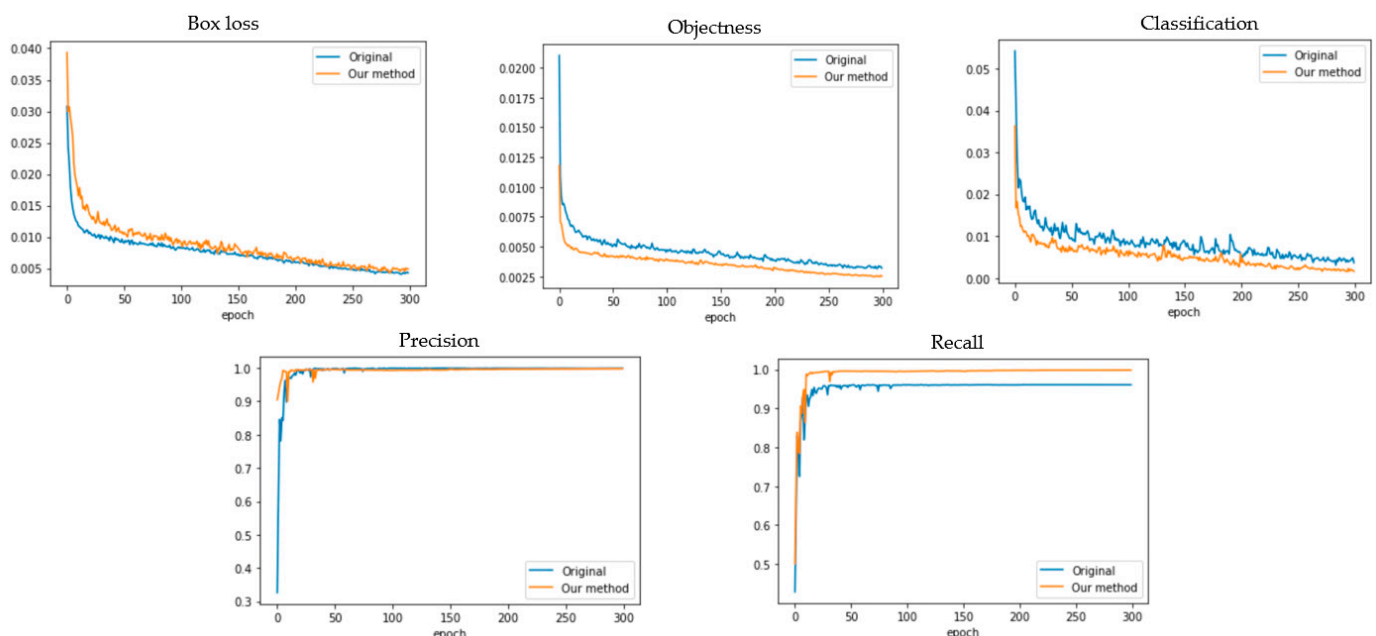


Figure 6. Training results of precious and proposed methods compared with Box loss, Objectness loss, Classification loss, Precision, and Recall.

The above results show that Objectness loss and Classification loss have significant improvement by using a modified version of YOLO. The previous versions of YOLOv3 and YOLOv4 consist of a bigger weight size compared to the new proposed version. Our model weight size is almost the same small size as the YOLOv4-tiny version. Their overall performances are compared in Table 2 with different previous YOLO models.

Table 2. Performance comparison results of the YOLO models.

Algorithms	Input Size	Training Accuracy	Testing Accuracy	Weight Size	Iteration Numbers	Training Time
YOLOv3	608 × 608	89.1%	88.1%	236 MB	50,000	65 h
YOLOv3-tiny		90%	85%	33.7 MB		24 h
Our model		99.4%	98.2%	31.2 MB		32 h
YOLOv4		93.9%	90.1%	245 MB		75 h
YOLOv4-tiny		89.9%	85.5%	23 MB		29 h

Our new augmentation method helped to increase both the training and testing accuracy from the above results, and a small disease area was even detected and classified with the new custom YOLOv4-tiny architecture. Figure 7 shows the detection and classification result with the test image.

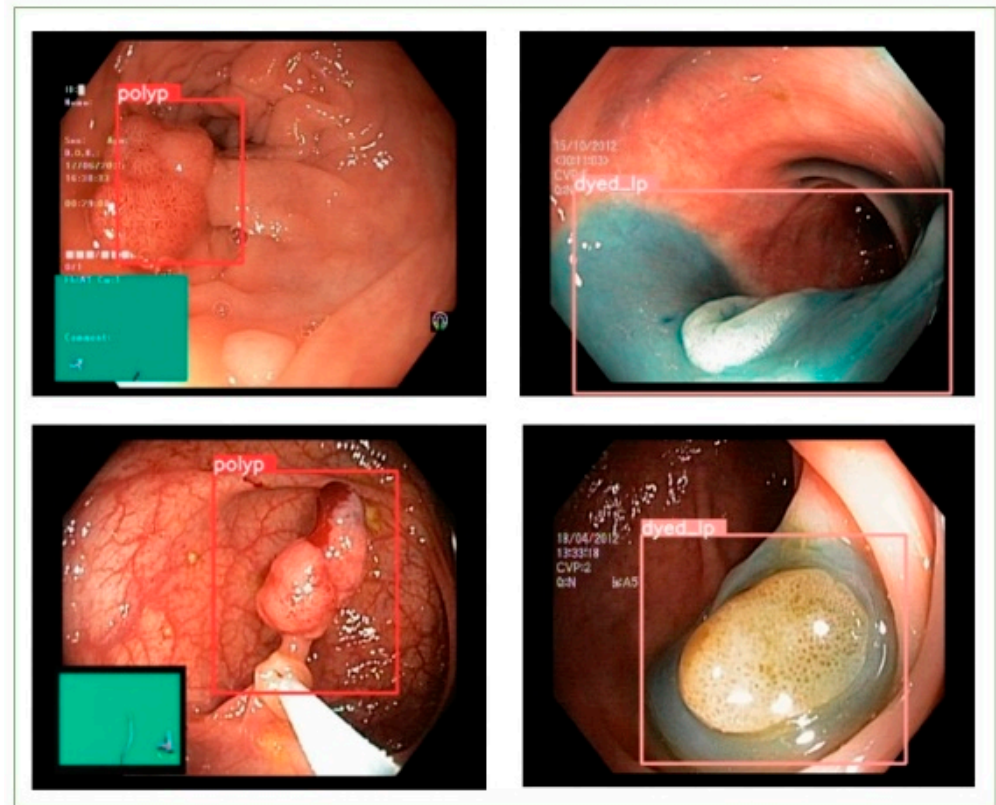


Figure 7. The detection and classification of the test image with our custom version.

6. Conclusions

In this study, in order to enhance the accuracy and speed of a custom classification and detection model, this research proposed a modified version of the YOLOv4-tiny model with the Inception-ResNet-A block for wireless. The study compared the previous original model architecture with a new custom version, and the result showed that under the same training and testing environment, the new custom version obtained a high positive performance in terms of detection and classification accuracy. By comparing the experimental result, loss of Objectness and loss of Classification are decreased by 1.75% and 2.03%, respectively. In addition, some previous data augmentation approaches effected a decrease in the efficiency of medical images because of sensitivity. Our new approach was achieved to give a limitation point in terms of color change. The proposed augmentation method did not face an overfitting problem while training and testing the wireless endoscopic images. In addition, the experimental result illustrates that the modified version of YOLOv4 achieved 99.4% training accuracy and 98.2% test accuracy, as compared with the original version of YOLOv4-tiny, which achieved 93.9% and 90.1%, respectively.

Wireless endoscopic image analysis for the detection of gastrointestinal tract diseases is a challenging and complicated task. This analysis task can be categorized into the classification and detection of images. The inconsistencies in endoscopic images occur due to organ movement and noise while recording the endoscopic images. Training a model on a small image dataset does not result in high accuracy. For this reason, we proposed a custom model.

In this study, we implemented data augmentation to increase the size of our dataset. To improve the proposed architecture, transfer learning or explainable artificial intelligence could also be considered.

Author Contributions: This research was designed and written by M.D.; conceptualization, M.D. and R.M.; methodology, M.D. and M.S.; software, M.D.; writing—original draft preparation, M.D.; supervised this study and contributed, Y.-I.C. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by Korea Agency for Technology and Standards in 2022, project numbers are K_G012002073401, K_G012002234001 and by the Gachon University research fund of 2020(GCU-202008460006).

Institutional Review Board Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Duan, S.; Gao, X.; Xia, C.; Ge, B. A2TPNet: Alternate Steered Attention and Trapezoidal Pyramid Fusion Network for RGB-D Salient Object Detection. *Electronics* **2022**, *11*, 1968. [\[CrossRef\]](#)
2. Habibzadeh, M.; Jannesari, M.; Rezaei, Z.; Baharvand, H.; Totonchi, M. Automatic white blood cell classification using pre-trained deep learning models: Resnet and inception. *Proc. SPIE* **2018**, *10696*, 1069612.
3. Rahim, T.; Usman, M.A.; Shin, S.Y. A survey on the contemporary computer-aided tumor, polyp, and ulcer detection methods in wireless capsule endoscopy imaging. *Comput. Med. Imag. Graph.* **2020**, *85*, 101767. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Deding, U.; Kaalby, L.; Bøggild, H.; Plantener, E.; Wollesen, M.K.; Kobaek-Larsen, M.; Hansen, S.J.; Baatrup, G. Colon Capsule Endoscopy vs. CT Colonography Following Incomplete Colonoscopy: A Systematic Review with Meta-Analysis. *Cancers* **2020**, *12*, 3367. [\[CrossRef\]](#)
5. Zimmermann-Fraedrich, K.; Sehner, S. Right-sided location not associated with missed colorectal adenomas in an individual-level reanalysis of tandem colonoscopy studies. *Gastroenterology* **2019**, *157*, 660–671. [\[CrossRef\]](#)
6. Chetcuti, S.; Sidhu, R. Capsule endoscopy-recent developments and future directions. *Expert Rev. Gastroenterol. Hepatol.* **2021**, *15*, 127–137. [\[CrossRef\]](#)
7. Sindhu, C.; Valsan, V. A novel method for automatic detection of inflammatory bowel diseases in wireless capsule endoscopy images. In Proceedings of the 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN), Chennai, India, 16–18 March 2017; pp. 1–6.
8. Alaskar, H.; Hussain, A.; Al-Aseem, N.; Liatsis, P.; Al-Jumeily, D. Application of Convolutional Neural Networks for Automated Ulcer Detection in Wireless Capsule Endoscopy Images. *Sensors* **2019**, *19*, 1265. [\[CrossRef\]](#)
9. Razzak, M.I.; Naz, S.; Zaib, A. Deep Learning for Medical Image Processing: Overview, Challenges and the Future. In *Classification in BioApps*; Dey, N., Ashour, A., Borra, S., Eds.; Lecture Notes in Computational Vision and Biomechanics; Springer: Cham, Switzerland, 2018; Volume 26. [\[CrossRef\]](#)
10. Ghosh, T.; Chakareski, J. Deep Transfer Learning for Automated Intestinal Bleeding Detection in Capsule Endoscopy Imaging. *J. Digit. Imaging* **2021**, *34*, 404–417. [\[CrossRef\]](#)
11. Tofanelli, M.B.D.; Wortman, S.E. Benchmarking the Agronomic Performance of Biodegradable Mulches against Polyethylene Mulch Film: A Meta-Analysis. *Agronomy* **2020**, *10*, 1618. [\[CrossRef\]](#)
12. Wei, X.; Wan, H.; Ye, F.; Min, W. HPS-Net: Multi-Task Network for Medical Image Segmentation with Predictable Performance. *Symmetry* **2021**, *13*, 2107. [\[CrossRef\]](#)
13. Weng, L.; Xu, Y.; Xia, M.; Zhang, Y.; Liu, J.; Xu, Y. Water Areas Segmentation from Remote Sensing Images Using a Separable Residual SegNet Network. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 256. [\[CrossRef\]](#)
14. Wang, C.; Zhang, R.; Chang, L. A Study on the Dynamic Effects and Ecological Stress of Eco-Environment in the Headwaters of the Yangtze River Based on Improved DeepLab V3+ Network. *Remote Sens.* **2022**, *14*, 2225. [\[CrossRef\]](#)
15. Bote-Curiel, L.; Muñoz-Romero, S.; Gerrero-Currieses, A.; Rojo-Álvarez, J.L. Deep Learning and Big Data in Healthcare: A Double Review for Critical Beginners. *Appl. Sci.* **2019**, *9*, 2331. [\[CrossRef\]](#)
16. Liu, P.; Song, Y.; Chai, M.; Han, Z.; Zhang, Y. Swin-UNet++: A Nested Swin Transformer Architecture for Location Identification and Morphology Segmentation of Dimples on 2.25Cr1Mo0.25V Fractured Surface. *Materials* **2021**, *14*, 7504. [\[CrossRef\]](#)
17. Ravanelli, M.; Paderno, A.; Del Bon, F.; Montalto, N.; Pessina, C.; Battocchio, S.; Farina, D.; Nicolai, P.; Maroldi, R.; Piazza, C. Prediction of Posterior Paraglottic Space and Cricoid Unit Involvement in Endoscopically T3 Glottic Cancer with Arytenoid Fixation by Magnetic Resonance with Surface Coils. *Cancers* **2019**, *11*, 67. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Safarov, S.; Whangbo, T.K. A-DenseUNet: Adaptive Densely Connected UNet for Polyp Segmentation in Colonoscopy Images with Atrous Convolution. *Sensors* **2021**, *21*, 1441. [\[CrossRef\]](#)
19. Inbaraj, X.A.; Villavicencio, C.; Macrohon, J.J.; Jeng, J.-H.; Hsieh, J.-G. Object Identification and Localization Using Grad-CAM++ with Mask Regional Convolution Neural Network. *Electronics* **2021**, *10*, 1541. [\[CrossRef\]](#)

20. Almotairi, S.; Kareem, G.; Aouf, M.; Almutairi, B.; Salem, M.A.-M. Liver Tumor Segmentation in CT Scans Using Modified SegNet. *Sensors* **2020**, *20*, 1516. [CrossRef]
21. Kvasir-Capsule, A Video Capsule Endoscopic Dataset. Scientific Data 2021. Available online: <https://datasets.simula.no/kvasir/> (accessed on 30 September 2022).
22. Prasath, V.B.S. Polyp Detection and Segmentation from Video Capsule Endoscopy: A Review. *J. Imaging* **2017**, *3*, 1. [CrossRef]
23. Li, G.; Sun, C.; Xu, C.; Zheng, Y.; Wang, K. Cervical Cell Segmentation Method Based on Global Dependency and Local Attention. *Appl. Sci.* **2022**, *12*, 7742. [CrossRef]
24. Visaggi, P.; Barberio, B.; Ghisa, M.; Ribolsi, M.; Savarino, V.; Fassan, M.; Valmasoni, M.; Marchi, S.; de Bortoli, N.; Savarino, E. Modern Diagnosis of Early Esophageal Cancer: From Blood Biomarkers to Advanced Endoscopy and Artificial Intelligence. *Cancers* **2021**, *13*, 3162. [CrossRef] [PubMed]
25. Barki, C.; Rahmouni, H.B.; Labidi, S. Prediction of Bladder Cancer Treatment Side Effects Using an Ontology-Based Reasoning for Enhanced Patient Health Safety. *Informatics* **2021**, *8*, 55. [CrossRef]
26. Grasso, S.M.; Peña, E.D.; Kazemi, N.; Mirzapour, H.; Neupane, R.; Bonakdarpour, B.; Gorno-Tempini, M.L.; Henry, M.L. Treatment for Anomia in Bilingual Speakers with Progressive Aphasia. *Brain Sci.* **2021**, *11*, 1371. [CrossRef] [PubMed]
27. Shakhnoza, M.; Sabina, U.; Cho, I.Y. Novel Video Surveillance-Based Fire and Smoke Classification Using Attentional Feature Map in Capsule Networks. *Sensors* **2021**, *22*, 98. [CrossRef] [PubMed]
28. Hoang, M.C.; Nguyen, K.T.; Kim, J.; Park, J.-O.; Kim, C.-S. Automated Bowel Polyp Detection Based on Actively Controlled Capsule Endoscopy: Feasibility Study. *Diagnostics* **2021**, *11*, 1878. [CrossRef]
29. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271. Available online: <https://arxiv.org/abs/1612.08242> (accessed on 22 November 2019).
30. Mokhtari, S.; Abbaspour, A.; Yen, K.K.; Sargolzaei, A. A Machine Learning Approach for Anomaly Detection in Industrial Control Systems Based on Measurement Data. *Electronics* **2021**, *10*, 407. [CrossRef]
31. Samir, J.; Ayan, S. A deep CNN model for anomaly detection and localization in wireless capsule endoscopy images. *Comput. Biol. Med.* **2021**, *137*, 104789.
32. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
33. Redmon, J.; Farhadi, A. Yolov3: An Incremental Improvement. 2018. Available online: <https://arxiv.org/abs/1804.02767> (accessed on 29 September 2022).
34. Szegedy, C.; Ioffe, S.; Vanhoucke, V. Inception-v4, Inception-Resnet and the Impact of Residual Connections on Learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
35. Bhatt, D.; Patel, C.; Talsania, H.; Patel, J.; Vaghela, R.; Pandya, S.; Modi, K.; Ghayvat, H. CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope. *Electronics* **2021**, *10*, 2470. [CrossRef]
36. Habijan, M.; Galić, I.; Romić, K.; Leventić, H. AB-ResUNet+: Improving Multiple Cardiovascular Structure Segmentation from Computed Tomography Angiography Images. *Appl. Sci.* **2022**, *12*, 3024. [CrossRef]
37. Fu, G.-H.; Wang, J.-B.; Zong, M.-J.; Yi, L.-Z. Feature Ranking and Screening for Class-Imbalanced Metabolomics Data Based on Rank Aggregation Coupled with Re-Balance. *Metabolites* **2021**, *11*, 389. [CrossRef]