Original Article

Introducing a Deep Neural Network Model with Practical Implementation for Polyp Detection in Colonoscopy Videos

Hajar Keshavarz[1], Zohreh Ansari[2], Hossein Abootalebian[1], Babak Sabet[3], Mohammadreza Momenzadeh[1]

[1]Department of Artificial Intelligence in Medical Sciences, Smart University of Medical Sciences, Tehran, Iran, [3]Department of Surgery, Faculty of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [2]Department of Biomedical Engineering, Engineering Faculty, Meybod University, Meybod, Yazd, Iran

Abstract

Background: Deep learning has gained much attention in computer-assisted minimally invasive surgery in recent years. The application of deep-learning algorithms in colonoscopy can be divided into four main categories: surgical image analysis, surgical operations analysis, evaluation of surgical skills, and surgical automation. Analysis of surgical images by deep learning can be one of the main solutions for early detection of gastrointestinal lesions and for taking appropriate actions to treat cancer. **Method:** This study investigates a simple and accurate deep-learning model for polyp detection. We address the challenge of limited labeled data through transfer learning and employ multi-task learning to achieve both polyp classification and bounding box detection tasks. Considering the appropriate weight for each task in the total cost function is crucial in achieving the best results. Due to the lack of datasets with nonpolyp images, data collection was carried out. The proposed deep neural network structure was implemented on KVASIR-SEG and CVC-CLINIC datasets as polyp images in addition to the nonpolyp images extracted from the LDPolyp videos dataset. **Results:** The proposed model demonstrated high accuracy, achieving 100% in polyp/non-polyp classification and 86% in bounding box detection. It also showed fast processing times (0.01 seconds), making it suitable for real-time clinical applications. **Conclusion:** The developed deep-learning model offers an efficient, accurate, and cost-effective solution for real-time polyp detection in colonoscopy. Its performance on benchmark datasets confirms its potential for clinical deployment, aiding in early cancer diagnosis and treatment.

Keywords: *Automatic polyp detection, deep learning, image processing, transfer learning*

Submitted: 07-Apr-2024 **Revised:** 04-Aug-2024 **Accepted:** 28-Nov-2024 **Published:** 09-Jun-2025

Introduction

Improving surveillance endoscopy, including colonoscopy, to prevent gastrointestinal cancers such as colorectal cancer, has always been an important area of research in gastrointestinal endoscopy. Detecting and removing adenomatous polyps protects patients from cancer progression. Therefore, the detection rate of polyps is considered an appropriate criterion for the quality of endoscopy.[1] Increasing the detection rate of polyps reduces mortality from colorectal cancer. However, the unacceptable high rate of missed neoplastic lesions still exists and varies among different endoscopists.[2] The wide range of size, tissue, shape, and color changes in polyps and the difficulty of distinguishing polyps from other tissues during endoscopy make this problem more complicated.[3] Thus, computer-aided polyp detection (CADe) systems were introduced to reduce the rate of polyp detection errors. Initial studies focused on the color and texture of polyps as the features that relevant specialists considered for each polyp.[4] Recently, newer methods based on deep neural networks (DNNs) for polyp detection have gained attention. The use of artificial intelligence-based CADe systems using DNNs for polyp detection has increased the detection rate of polyps in various studies.[5] Many commercial CADe systems have been introduced worldwide, including GI Genius,[6] Discovery,[7] EndoMind,[8] Ai4Gi,[9] and EndoBrain,[10] which utilize artificial intelligence for this purpose.

In an ideal situation, CADe systems should have constant performance and robust output against images from different patients, using different devices and endoscopy procedures, and detect or segment polyps in real time. To achieve appropriate performance of DNN structures, extensive data collection and labeling

How to cite this article: Keshavarz H, Ansari Z, Abootalebian H, Sabet B, Momenzadeh M. Introducing a deep neural network model with practical implementation for polyp detection in colonoscopy videos. J Med Signals Sens 2025;15:17.

Address for correspondence:
Dr. Mohammadreza Momenzadeh,
No. 3, 1st Alley, Sarafraz St., Shaheed Beheshti St., Tehran, Iran.
E-mail: momenzadeh.mr@gmail.com

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.jmss_23_24

Quick Response Code:

by relevant specialists are required. Transfer learning is implemented as a method to overcome the problem of insufficient training data for deep learning systems.[11,12] This concept refers to using pretrained deep convolutional neural networks (CNNs) on large natural image datasets, which have been widely used in various applications in medicine.

In this article, a DNN structure for polyp detection in endoscopic images is introduced that can be implemented with low memory and utilizes transfer learning to overcome the problem of insufficient training data. The novelty of this article is proposing an optimal simple and accurate model for polyp detection instead of using complicated structures such as R-CNN, Retina Net, Dark Net, and Yolo with so many parameters to be trained. Another novelty of this work is considering a weighted loss function to fine-tune this network to extract features for both object detection and classification. Our study contributes to this domain by proposing a neural network structure that integrates transfer learning and multi-task learning to handle both classification and detection tasks simultaneously.

We organized this paper as follows: the common DNN structures implemented in the polyp detection studies are introduced in section two. In section three, the proposed DNN structure is described. Furthermore, in section four, the experimental results obtained from the utilization of KVASIR-SEG and CVC-CLINIC datasets are presented,[13,14] and in section five, conclusions are provided.

Related Works

Many studies have aimed to enhance the detection of polyps during colonoscopy using CNNs. In a study by Pozdeev *et al*.,[15] a fully convolutional network (FCN) with a novel structure that initially makes binary classification predictions and then processes them through a CNN similar to the U-Net architecture is discussed. The proposed network achieves state-of-the-art performance on the Kvasir-SEG and CVC-ClinicDB datasets. In a study by Lequan Yu *et al*.,[16] a three-dimensional, FCN for segmentation has been proposed. PraNet[17] provided real-time segmentation capabilities using deep supervision mechanisms and a reverse attention module for boundary detection. PraNet also features a parallel partial decoder to improve its performance, and these modules have been incorporated into other innovative architectures like AMNet.[18] Furthermore, a study by Tomar *et al*.[19] introduced a dual-decoder network called DDANet, which employs a single ResNet-style encoder with a dual-decoder architecture to generate both a grayscale image and a segmentation mask.

In a study by Shin *et al*.,[3] a region-based CNN structure was used for automatic polyp detection. In this article, the authors used an Inception ResNet structure for polyp detection on colonoscopy images, utilizing transfer learning and postprocessing methods. In a study by Lee *et al*.,[20] the YOLOv2 structure was explored to develop algorithms for polyp detection and localization. Further studies were done[21] to use improved YOLO network structures for polyp detection. The YOLO structures, as shown in Figure 1, consist of input, backbone, neck, and head. The Backbone section, which performs feature extraction from images, includes different CNN structures such as VGG16, ResNet50, ResNet101, and DarkNet. The neck, which extracts information from the layers of different networks placed in the backbone, can be FPN, PANet, and BiFPN networks. The structures located in the neck can be used to extract information related to different objects of different sizes. Finally, in the head section, detection networks such as RCNN and RFCN structures are used to detect bounding boxes.

In a study by Nogueira-Rodríguez *et al*.,[22] a deep learning model for real-time polyp detection based on a pretrained YOLOv3 architecture and a postprocessing step based on an object-tracking algorithm to reduce false positives (FPs) is proposed. In YOLOv3, on top of Darknet53, 53 more layers were added for proper object detection, making it a 106-layer CNN. The research done by Jha *et al*.[21] provided a benchmark for polyp detection, localization, and colonoscopy segmentation on recent state-of-the-art deep learning algorithms, including FasterRCNN, RetinaNet, YOLOv3 + SPP, YOLOv4, and EfficientDet. It compared the execution performances and the differences between various algorithms on variable polyp sizes and image resolutions. In this article, the ColonSegNet is proposed to achieve a better trade-off between an average precision and mean Intersection over Union (IoU) and the fastest detection and localization task. One of the models used in the study was EfficientDet, which is based on EfficientNet backbone architecture. Other models used were Faster R-CNN, YOLOv3, and YOLOv4. Qian *et al*.[23] also presented an enhanced FasterRCNN-based system for polyp detection.

Reddy *et al*.[24] utilized the advanced YOLO-v4 detection model for polyp detection during colonoscopy. The model's performance was evaluated based on the Multiple Object Tracking Accuracy and Multiple Object Tracking Precision criteria of the custom dataset. In a study by Ghose *et al*.,[25] a fine-tuned YOLO-v5 model for polyp colonoscopy was proposed. The model achieved competitive results with other deep learning techniques such as R-CNN, Faster R-CNN, and YOLO-v4. Three variants of YOLOv5 (s, m, and l) were employed by Al Amin *et al*.[26] to analyze Kvasir-SEG data. According to the research, YOLOv5l outperformed the other models in polyp identification, with an IoU of 86.25%.

All the indicated methods employ complex architectures requiring substantial computational resources. In contrast, our method emphasizes a balance between computational

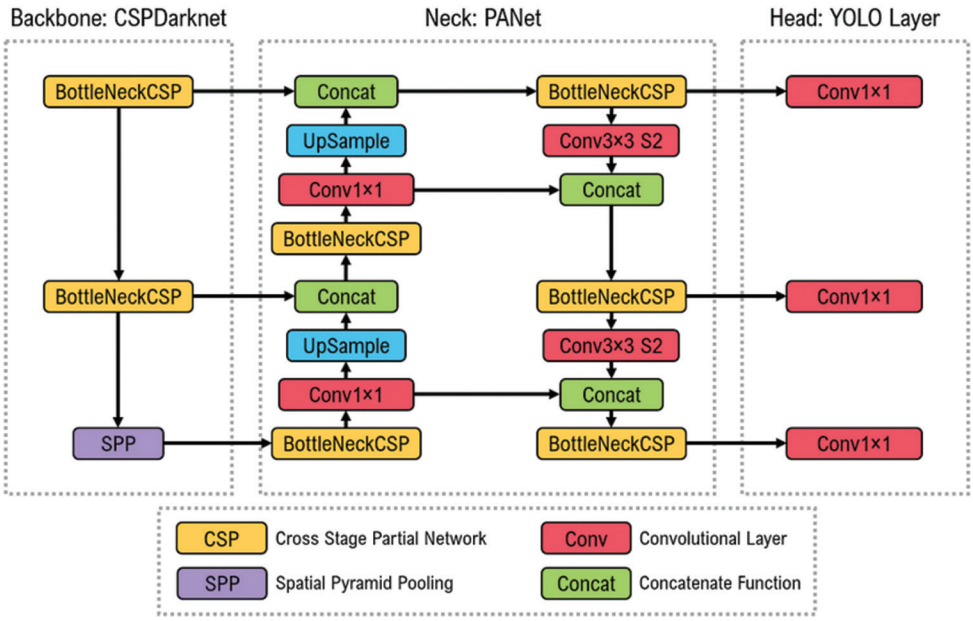

Figure 1: YOLO structure

efficiency and detection performance. By utilizing a pretrained ResNet50 for feature extraction and a streamlined network for detection and classification, our model is both resource-efficient and highly accurate. This makes our approach more accessible for deployment in resource-limited settings without sacrificing performance.

The ENDOMIND software, which is a publicly available software, is based on the YOLOv5 structure. The weights of the YOLOv5 structure used in this study were pretrained on public colonoscopy datasets and then fine-tuned on datasets collected from German hospitals. However, due to the complexity of the structure, its usage is usually limited to fine-tuning its parameters on pretrained networks for the desired dataset, and modification of the structure is challenging. Another problem with using this structure is its requirement for graphics processing unit (GPU) for computations. Therefore, we designed and trained our neural network structure and evaluated its results to overcome these challenges and improve the structure.

The Proposed Neural Network Structure

Our designed DNN structure consists of feature extraction, regression, and classification modules [Figure 2]. The pretrained ResNet50 CNN, which was trained on more than one million images on the ImageNet dataset, was used for transfer learning to extract features from input images. This structure comprises a 50-layer CNN with residual connections. In this article, ResNet 50 is considered as the first module for feature extraction and simply, the regression and classifier modules are considered on top of it. As indicated in a study by Ahamed *et al*.,[27] ResNet50, with distinctive residual connections and 50-layer depth, efficiently tackles the complexities of training DNNs. Therefore, this model serves as a strong basis for building specialized models with improved generalization abilities Moreover, this model, on one hand, has a small number of parameters concerning other deep structures, making the model not overfit on the training data, with respect to Resnet 101 and Darknet 53[28] and, on the other hand, achieves comparable polyp detection results with respect to the proposed Yolo structures as it is shown in Table 1.

The regression and classification modules are designed by three-layer feedforward neural network structures that receive the extracted features from the last hidden layer of the ResNet network as input and calculate polyp bounding boxes information in the output after several processing steps in the regression section. The presence or absence of polyps in each image frame is detected in the classification section.

To train this network structure, the mean squared error (MSE) loss function was used,[29] which is calculated based on the weighted sum of the MSE loss function for the regression module and defined as follow:

$$\text{Total MSE} = \lambda\left(\sum_{i=1:N}\sum_{j=1:M}(d_1^{ij}-z_1^{ij})^2\right) + (1-\lambda)\left(\sum_{i=1:N}\sum_{j=1:M}(d_2^{ij}-z_2^{ij})^2\right) \quad (1)$$

where, is the desired output of each neuron j in the regression module for input i, and is the calculated output by the network for that neuron. Similarly, and are the desired and calculated output of neuron j in the classification section for each input I, respectively. The coefficient λ represents the importance of the regression and classification errors in training the entire network weight matrix, which is determined during the training process.

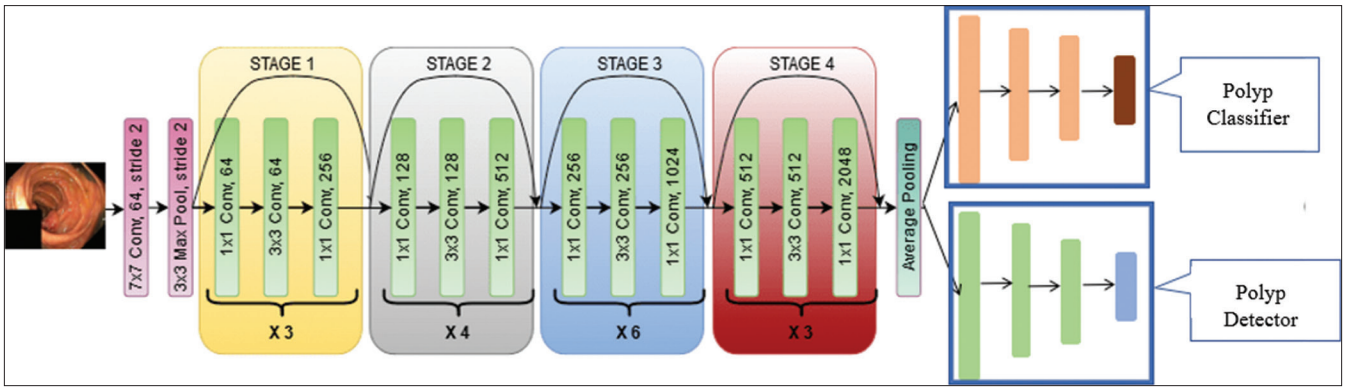

Figure 2: Proposed deep neural network structure, consisting of three parts: Feature extraction, classifier, and regression modules

Table 1: Comparison of accuracy and processing time of the proposed structure with the EndoMind software

Used structure	TP	Time processing (s)
EndoMind	88	0.02
Proposed structure	86	0.01

TP – True positive

Experimental Results

Experimental setup

Initially, the KVASIR-SEG dataset was implemented for training the network. This dataset contains the largest number of labeled samples, including images of polyp masks as well as their bounding box information in a json file.

It comprises 1000 colonoscopy images with polyps and their corresponding masks. The resolution of images in this dataset varies from 487 × 332 to 1072 × 1920 pixels. To increase the training data, the CVC-Clinc dataset images were also added to the KVASIR dataset. This dataset consists of 612 colonoscopy images with polyps and their mask images.

Due to the lack of datasets without polyps, noise-free frames with no blurring or specularity were selected from the videos in the LDPolyp dataset. Among these frames, 1586 images were selected for training the network. Then, to have the same dimensions for all images, the input image size was changed to 224 × 224 pixels. Some examples of nonpolyp images used for training the proposed network are shown in Figure 3.

The performance evaluation of the proposed DNN structure for object detection was conducted by measuring the overlap range of detected bounding boxes (A) by the network, with the labeled bounding boxes (B) using IoU, which is calculated as shown in eq. 2.[30]

$$\mathrm{IoU}(A, B) = \frac{A \cap B}{A \cup B} \quad (2)$$

Experiments were conducted using the PyTorch library on a system with the following specifications: Intel Core i7 6700 3.4 GHz, 32 GB RAM, Nvidia Geforce GTX1060 6GB.

Results

The experiments reported in this section include (1) evaluation of the proposed structure on KVASIR and CVC-CLINIC datasets. Since these datasets include only polyp images, the implemented structure includes only a regression module, (2) Addition of generated nonpolyp images to the previous dataset and training of the proposed structure with both classifier and regression modules. And, (3) Comparing the performance of the proposed structure with Yolov5 as the state-of-the-art model in object detection.

As the first step, the bounding box information of the polyps was extracted from the available mask of each image in the datasets and used as the label for training the network. Then, in the first experiment, the proposed DNN structure (without the classification branch) was trained on the KVASIR-SEG dataset. Eighty percent of the images were considered for training and the remaining 20% were used to test the model. The data were divided into batches of 32 samples for training the model. The MSE loss function and Adam optimizer were used to adjust the parameters. The learning rate of the model was set to 0.001, and the network was trained for 20 iterations. The achieved results of the network on the test set are reported in Table 2. By adding the CVC-CLINIC dataset to the training data, the performance of the network was improved as shown in Table 2. In this table, the number of test images with the IoU > 0.5 is considered true positive (TP), the number of test images with an IoU less than this value is considered FP, and false negative (FN) is the number of test images with IoU equal to zero.

As shown in the above table, increasing the diversity of the training data by combining two datasets has improved the network's generalizability and improved its performance by up to 7%. Figure 4 shows some examples of the proposed neural network's performance in polyp detection. As observed in this figure, the designed DNN is capable of accurately detecting polyp bounding boxes even with missing visual information.

In addition to the above cases, this network has been able to accurately predict the presence of hidden polyps in images that have polyps but are not correctly labeled by

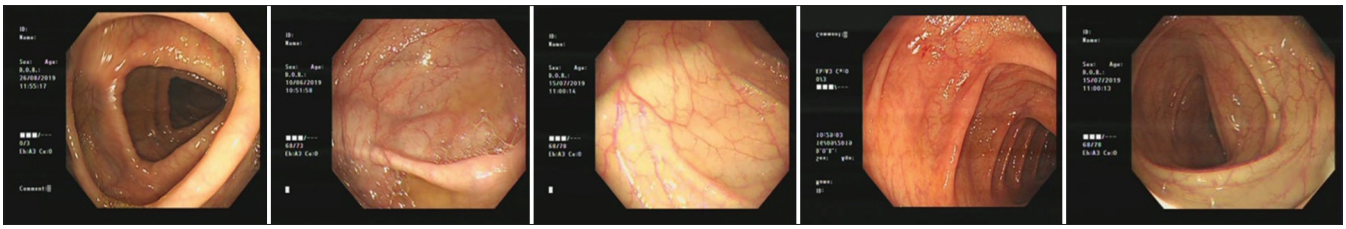

Figure 3: Some sample images collected from the nonpolyp dataset

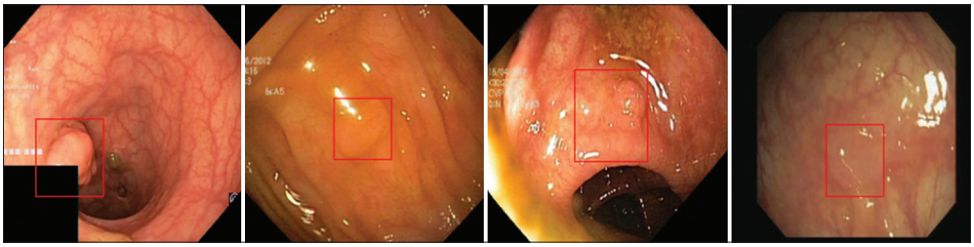

Figure 4: Some examples of the proposed network's performance in detecting polyp bounding boxes. (The detected polyps are shown by red squares)

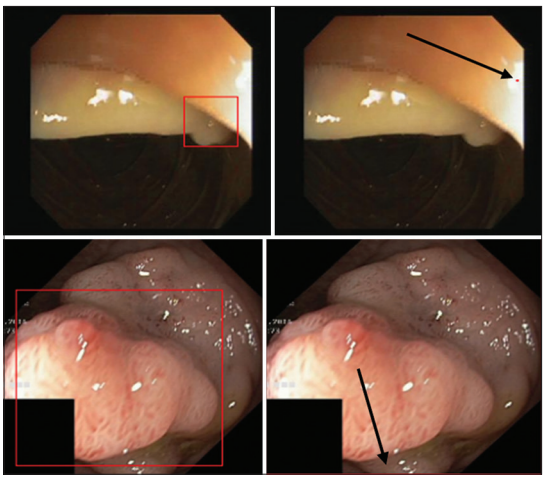

Figure 5: Polyp detection by the proposed structure (right column) for images with polyps that have not been labeled correctly by the specialist. The specialist labels are shown by arrows (left column)

Table 2: Performance of the network using the Kvasir-SEG dataset and the combined Kvasir-SEG and CVC-CLINIC datasets as training data

Trained dataset	TP (%)	FP (%)	FN (%)
Kvasi-SEG	66	24	10
Kvasir-SEG + CVC-CLINIC	71	18	11

TP – True positive; FP – False positive; FN – False negative

Table 3: The confusion matrix calculated from evaluating the trained network on the polyp dataset

Actual	Predicted	
	Positive (%)	Negative (%)
Positive	TP=72.88	FP=17
Negative	FN=10	TN=UNK

TP – True positive; FP – False positive; FN – False negative; TN – True positive, UNK – Unknown

the specialists. Figure 5 shows some examples of Polyp detection by the proposed method that have not been labeled by the specialist.

Due to the limited availability of data, cross-validation was used for a more accurate evaluation. In this case, we used 5-fold cross-validation which randomly divided data into five folds, and each time, 4 folds were considered for training and 1-fold for testing. The confusion matrix by averaging the results on test images, considering different ranges of IoU between the detected bounding boxes and the labeled ones, is provided in Table 3.

In the above table, the true negative (TN) parameter is not defined due to the lack of data without polyps.

In the second experiment, the addition of data without polyps and considering 80% of the total data for training and the remaining 20% for testing while using stratified sampling to adjust the diversity of data with and without polyps in each batch of training and testing, the model was trained. In this step, the classification and the regression modules of the model were simultaneously trained using backpropagation to minimize the total loss function, which is calculated as the weighted sum of the loss function of each module.

The evaluation of the proposed model with both classifier and regression modules includes evaluating its object detection and classification performances. Therefore, Table 4 is considered to verify the object detection performance by calculating the confusion matrix based on the IoU parameter. Therefore, for polyp images, the number of test images with an IoU > 0.5 is considered TP, and those with an IoU less than this value are considered FN. For nonpolyp images, the number of test images with calculated bounding boxes for them is considered as FP, and those with no object prediction are considered as TN. In Table 5, we evaluated the confusion matrix for the classification task. Our experimental results show the TP and TN values of 100% and FP and FN values of zero.

Table 4: The confusion matrix calculated for evaluating the object detection of the trained network on the polyp and nonpolyp datasets

Actual	Predicted	
	Positive (%)	Negative (%)
Positive	TP=71.72	FP=0
Negative	FN=28.25	TN=100

TP – True positive; FP – False positive; FN – False negative; TN – True positive

Table 5: The confusion matrix calculated from evaluating the classifier module of the trained network on the polyp and nonpolyp datasets

Actual	Predicted	
	Positive (%)	Negative (%)
Positive	TP=100	FP=0
Negative	FN=0	TN=100

TP – True positive; FP – False positive; FN – False negative; TN – True positive

The required time for training the entire weight of this network structure using GPU is 15.02 s with a standard deviation of 0.13, and the required time for processing each image by this model is 0.13 s with a standard deviation of 0.007 s. It should be noted that this structure can also be implemented on a central processing unit.

The Yolov5 structure is considered in the EndoMind software to detect polyp bounding boxes. Therefore, in the third experiment, the performance of the proposed DNN is compared with the Yolo-v5 on KVASIR-SEG data, which is considered one of the training datasets for both structures. Since EndoMind has been trained on the whole KVASIR-SEG dataset in addition to the other datasets, the evaluation of both models was performed when our model was trained and tested on the total KVASIR-SEG dataset, too. In Table 1, you can find the comparison of the performance of these two structures as well as the processing time requirement of each structure per image.

Discussion and Conclusion

In this paper, a DNN structure for polyp detection on endoscopic images was introduced. The simplification of the network structure resulted in a decrease in the model test time to 0.01 s per frame, which is suitable for real-time applications. The proposed structure has low structural complexity for practical implementation in hardware with low cost. Due to the lack of labeled data of endoscopic videos, transfer learning was used in the pretraining of this structure. Furthermore, the network was trained for two tasks, namely detecting polyp bounding boxes and determining the presence/absence of polyps, by using multi-task learning and adjusting the weight of each task in the training process. In this regard, data collection was also carried out for images without polyps. The results verify that this structure was able to correctly detect polyps and ignore the labeling errors made by experts in some cases. In addition, by increasing the diversity of the training data, the generalization capability of the network improved. This result indicates the need for using more diverse data to improve the performance of the proposed structure. Moreover, to ensure the reliability of the results due to the lack of training data, cross-validation was used to evaluate the model by considering different parts of the data for training and testing. The confusion matrix obtained from the evaluation of the network with both regression and classification modules indicates that this network structure was also able to detect polyp bounding boxes in polyp frames with an acceptable accuracy while detecting all nonpolyp frames.

In conclusion, our proposed method offers several advantages, including a simplified architecture suitable for low-cost hardware, efficient processing times, and robust performance even with limited training data. By integrating transfer learning and multi-task learning, our model provides a comprehensive solution for real-time polyp detection in endoscopic and colonoscopy videos. We believe that these comparisons with existing successful methods demonstrate the efficacy and practicality of our approach.

Financial support and sponsorship

This study is funded by Smart University of Medical Sciences, Tehran, Iran. Based on the agreement between Smart University of Medical Sciences and the editor of the journal, Dr. Rabbani, regarding the publication of the best articles of "the first International Congress of Artificial Intelligence in Medical Sciences" the paper titled 'Introducing a DNN structure with practical implementation capability for polyp detection in endoscopic and colonoscopy videos, and the agreement attached please find the final manuscript of the paper along with the supporting files.

Conflicts of interest

There are no conflicts of interest.

References

1. Kaminski MF, Robertson DJ, Senore C, Rex DK. Optimizing the quality of colorectal cancer screening worldwide. Gastroenterology 2020;158:404-17.
2. Zhao S, Wang S, Pan P, Xia T, Chang X, Yang X, *et al.* Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: A systematic review and meta-analysis. Gastroenterology 2019;156:1661-74.e11.
3. Shin Y, Qadir HA, Aabakken L, Bergsland J, Balasingham I. Automatic colon polyp detection using region based deep CNN and post learning approaches. IEEE Access 2018;6:40950-62.
4. Karkanis SA, Iakovidis DK, Maroulis DE, Karras DA, Tzivras M. Computer-aided tumor detection in endoscopic video using color wavelet features. IEEE Trans Inf Technol Biomed 2003;7:141-52.

5. Hassan C, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, *et al.* Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: A systematic review and meta-analysis. Gastrointest Endosc 2021;93:77-85.e6.
6. Repici A, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, *et al.* Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. Gastroenterology 2020;159:512-20.e7.
7. Hann A, Troya J, Fitting D. Current status and limitations of artificial intelligence in colonoscopy. United European Gastroenterol J 2021;9:527-33.
8. Fitting D, Krenzer A, Troya J, Banck M, Sudarevic B, Brand M, *et al*. A video based benchmark data set (ENDOTEST) to evaluate computer-aided polyp detection systems. Scand J Gastroenterol 2022;57:1397-403.
9. Byrne MF, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, *et al*. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. Gut 2019;68:94-100.
10. Mori Y, Kudo SE, Chiu PW, Singh R, Misawa M, Wakamura K, *et al*. Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: An international web-based study. Endoscopy 2016;48:1110-8.
11. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, *et al*. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Trans Med Imaging 2016;35:1299-312.
12. Escobar J, Sanchez K, Hinojosa C, Arguello H, Castillo S. Accurate Deep Learning-Based Gastrointestinal Disease Classification Via Transfer Learning Strategy. XXIII Symposium on Image, Signal Processing and Artificial Vision (STSIVA). IEEE; 2021.
13. Jha D, Smedsrud PH, Riegler MA, Halvorsen P, de Lange T, Johansen D, *et al*. Kvasir-Seg: A segmented polyp dataset. In: MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, Proceedings, Part II 26. Springer International Publishing; 2020. p. 451-62.
14. Bernal J, Sánchez FJ, Fernández-Esparrach G, Gil D, Rodríguez C, Vilariño F. WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation versus saliency maps from physicians. Comput Med Imaging Graph 2015;43:99-111.
15. Pozdeev AA, Obukhova NA, Motyko AA. Automatic Analysis of Endoscopic Images for Polyps Detection and Segmentation. In: 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). IEEE; 2019. p. 1216-20.
16. Lequan Y, Hao Ch, Qi D, Jing Q, Pheng AH. Integrating online and offline three-dimensional deep learning for automated polyp detection in colonoscopy videos. IEEE J Biomed Health Inform 2017;21:65-75.
17. Fan DP, Ji GP, Zhou T, Chen G, Fu H, Shen J, *et al*. Pranet: Parallel Reverse Attention Network For Polyp Segmentation. In: International Conference on Medical Image Computing and Computer-assisted Intervention. Berlin: Springer; 2020. p. 263-73.
18. Song P, Li J, Fan H. Attention based multi-scale parallel network for polyp segmentation. Comput Biol Med 2022;146:105476.
19. Tomar NK, Jha D, Ali S, Johansen HD, Johansen D, Riegler MA, *et al*. Ddanet: Dual Decoder Attention Network For Automatic Polyp Segmentation. In: Pattern Recognition. ICPR International Workshops and Challenges: Virtual Event.
20. Lee JY, Jeong J, Song EM, Ha C, Lee HJ, Koo JE, *et al*. Real-time detection of colon polyps during colonoscopy using deep learning: Systematic validation with four independent datasets. Sci Rep 2020;10:8379.
21. Jha D, Ali S, Tomar NK, Johansen HD, Johansen D, Rittscher J, *et al*. Real-time polyp detection, localization and segmentation in colonoscopy using deep learning. IEEE Access 2021;9:40496-510.
22. Nogueira-Rodríguez A, Domínguez-Carbajales R, Campos-Tato F, Herrero J, Puga M, Remedios D, *et al*. Real-time polyp detection model using convolutional neural networks. Neural Comput Appl 2022;34:10375-96.
23. Qian Z, Lv Y, Lv D, Gu H, Wang K, Zhang W, *et al*. Anew approach to polyp detection by pre processing of images and enhanced faster R-CNN. IEEE Sensors J 2021;21:11374-81.
24. Reddy JS, Venkatesh C, Sinha S, Mazumdar S. Real Time Automatic Polyp Detection in White Light Endoscopy Videos Using a Combination of YOLO and DeepSORT. In: Proceeding 1st International Conference on the Paradigm Shifts in Communication, Embedded Systems, Machine Learning and Signal Processing (PCEMS); 2022. p. 104-6. [doi: 10.1109/PCEMS55161.2022.9807988].
25. Ghose P, Ghose A, Sadhukhan D, Pal S, Mitra M. Improved polyp detection from colonoscopy images using finetuned YOLO-v5. Multimed Tools Appl 2023;83:42929-54.
26. Al Amin M, Paul BK, Bithi NI. Real Time Detection and Localization of Colorectal Polyps from Colonoscopy Images: A Deep Learning Approach. In: Proceeding 2022 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE); 2022. p. 58-61. [doi: 10.1109/WIECONECE57977.2022.10151209].
27. Ahamed MF, Islam MR, Nahiduzzaman M, Karim MJ, Ayari MA, Khandakar A. Automated detection of colorectal polyp utilizing deep learning methods with explainable AI. IEEE Access 2024;12:78074-100.
28. Li Z, Peng C, Yu G, Zhang X, Deng Y, Sun J. Detnet: Design Backbone for Object Detection. In: Proceedings of the European Conference on Computer Vision (ECCV); 2018. p. 334-50.
29. Christoffersen P, Kris J. The importance of the loss function in option valuation. J Financ Econ 2004;72:291-318.
30. Rahim T, Ali Hassan S, Shin SY. A deep convolutional neural network for the detection of polyps in colonoscopy images. Biomed Signal Process Control 2021;68:102654.