ISSN 1870-4069

Identification and Segmentation of Polyps in the Colon Applying Artificial Intelligence Techniques: A Systematic Literature Review

Valentina Cardenas Moreno[1], Luz Ivana Correa Hernández[2], Carlos Alberto López Herrera[3], Héctor Gabriel Acosta Mesa[3], Efrén Mezura Montés[3]

[1] Universidad Veracruzana, Facultad de Instrumentación Electrónica, Mexico

[2] Universidad Veracruzana, Facultad de Administración y Contaduría, Mexico

[3] Universidad Veracruzana, Instituto de Investigaciones en Inteligencia Artificial, Mexico

emezura@uv.mx

Abstract. Colorectal cancer, the fourth most common type of cancer worldwide, can be prevented through the early identification of colorectal polyps. In this regard, artificial intelligence has emerged as a promising tool to improve the identification, segmentation, and classification of polyps in medical images. The application of these techniques has been shown to aid medical diagnosis due to their accuracy and effectiveness in detecting polyps. This paper reviews the state of the art of existing artificial intelligence techniques, the most common architectures, the datasets used for training, and the metrics employed to evaluate the performance of the proposed models. Finally, opportunities for improvement are highlighted, and future research directions are proposed to optimize artificial intelligence-assisted diagnosis in gastrointestinal health.

Keywords: Colorectal polyps, artificial intelligence, datasets, performance metrics.

1 Introduction

Colon cancer is the fourth most common type of cancer worldwide and the third most common cause of cancer-related deaths [1]. This type of cancer typically develops from untreated colorectal polyps, which are caused by the growth of mucosal epithelial cells [2] and progress slowly. If not properly identified and treated within a period of 10 to 20 years, adenocarcinomatous polyps can lead to colorectal cancer [1].

The timely and early detection of colorectal polyps is essential for the prevention of colon cancer [3]. Therefore, it is crucial to minimize the occurrence of false negatives during the analysis of a colonoscopy study, a medical procedure in which an instrument is used to film and examine the interior of the colon [4].

In the modern context of Artificial Intelligence (AI), there is no standardized or widely adopted technique in medical practice. Therefore, this systematic review analyzes proposals aimed at implementing computer-aided diagnosis by applying techniques that facilitate the identification, segmentation, and/or classification of polyps for timely detection, thereby exploring more accessible and efficient alternatives.

2 Method

For the methodology of the systematic literature review (SLR), the model proposed by Kitchenham and Charters [5] was used, as it is specialized in computer science and provides a relevant and precise approach. Additionally, being a strict and well-structured model, it helps minimize biases in the selection of papers.

2.1 Justification

The timely detection of colorectal polyps is key to the prevention of colorectal cancer. Therefore, the aim of this review is to understand the state of the art of AI techniques applied to the identification and segmentation of intestinal polyps and how these support tools improve medical diagnosis.

2.2 Research Questions

The following questions were used to guide the research and narrow down the topic.

RQ1. What artificial intelligence techniques have been proposed for the identification and segmentation of intestinal polyps?

RQ2. What are the characteristics (name, type of access, number, format, resolution, and modality) of the image datasets available in the literature for the classification of intestinal polyps?

RQ3. What are the most commonly used performance metrics to evaluate popular AI techniques in the identification and segmentation of polyps?

2.3 Objectives

General objective. To investigate and understand the state of the art in the application of AI techniques for the identification and segmentation of intestinal polyps.

Specific objectives. 1. Identify the existing AI techniques used for the classification of intestinal polyps. 2. Recognize the characteristics of the image datasets available in the literature for the classification of intestinal polyps. 3. Understand the most commonly used performance measures to evaluate the most popular AI techniques.

Table 1. Inclusion and exclusion criteria applied to the papers retrieved with the search strings.

Inclusion criteria	Exclusion criteria
IC1 Studies published from 2019 to September 2024. **IC2** Open Access studies.	**EC1** Studies that are in a language other than English. **EC2** Studies that are not research papers.
IC3 Studies that contain at least two keywords in the title. **IC4** Studies that contain at least three keywords in the abstract.	**EC3** Studies whose title refers to polyps outside of the colon. **EC4** Studies conducted with proprietary datasets (or, alternatively, not publicly accessible datasets).
IC5 Studies that, when reading their abstract and/or conclusion, mention the AI technique applied, the datasets used, and the performance achieved.	

2.4 Search

Keywords selection. Based on the research questions and objectives, the necessary keywords were defined, along with their synonyms, as follows: “intestinal polyps,” “colon polyps,” “polyps,” “artificial intelligence,” “AI,” “machine learning,” “performance,” “techniques,” “classification,” and “dataset.”

Search strings. The following academic databases were selected: (1) IEEE Xplore Digital Library, (2) Wiley Online Library, (3) SpringerLink, (4) ACM Digital Library, and (5) ScienceDirect, where search strings were tested to finally select two strings, with one being exclusive for ScienceDirect due to its limitation of only supporting eight boolean operators.

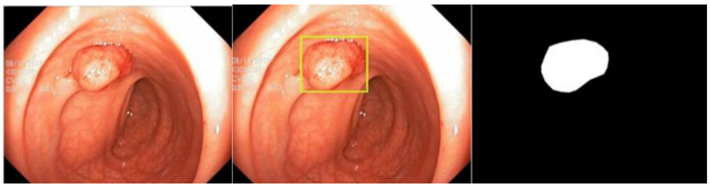
- Search string for IEEE, Wiley, Springer Link and ACM: (“intestinal polyps” OR “polyps” OR “colon polyps”) AND (“artificial intelligence” OR “AI” OR “machine learning”) AND (“classification” OR “techniques” OR “performance” OR “dataset”).
- Search string for ScienceDirect: (“intestinal polyps” OR “polyps” OR “colon polyps”) AND (“artificial intelligence” OR “AI”) AND (“classification” OR “techniques” OR “performance” OR “dataset”).

Inclusion and exclusion criteria. The established criteria are shown below in Table 1.

2.5 Data Selection and Extraction

A total of 1288 papers were identified in the five selected academic databases. After applying the criteria IC1, IC2, EC1, and EC2, the sample was reduced to 161 papers. Subsequently, using the criteria IC3, IC4, IC5, EC3, and EC4, 26 papers were selected for analysis in section 3.

For data extraction, key information was collected such as the paper title, publication date, authors, technique or method used, task type (identification, segmentation, and/or classification), datasets employed, technique description, performance metrics, and



..(a) (b) (c)

Fig. 1. (a) Original image from a colonoscopy; (b) identification with bounding box; (c) segmentation with the generated segmentation mask. [7].

areas of opportunity. Please refer to section 4. Discussion for the link to access the retrieved data from each of the selected papers.

3 Results

In this section, the results found within the 26 selected papers are presented, and the research questions are answered.

3.1 Selected studies and their characteristics

Of the 26 selected papers, 13 belong to IEEE, 7 to Wiley, and 6 to SpringerLink, with publications between 2019 and September 2024. Most of these studies focus on the use of artificial intelligence for medical image analysis to improve the detection of colorectal polyps. In particular, 20 papers explore methods related to the detection and delineation of polyps, while the remaining 6 focus on tasks related to the classification of gastrointestinal diseases.

All the proposed techniques employ artificial intelligence models trained with publicly or privately accessible datasets and are evaluated using various parameters known as performance metrics.

3.2 Identification, Segmentation and Classification

During the review, it was identified that the terms identification and segmentation can be used interchangeably in the medical field. However, in the context of AI techniques, these concepts have more specific differences that should be considered for a better understanding of their application. This was a key finding after the detailed analysis of the reviewed papers.

Identification of colorectal polyps refers to their detection within the image or video obtained during colonoscopy. In this process, the polyp is outlined with a rectangle indicating its location, also known as a "bounding box" [4]; as shown in Figure 1(b). In contrast, segmentation aims to produce a mask (segmentation mask) that separates the areas of interest (polyps) from the healthy areas (colon), allowing the delineation of the polyp's body and edges. As a result, shown in Figure 1(c), an image is obtained with the areas of interest in white and the rest in black [6].

On the other hand, classification can have two meanings. In the context of medical evaluation, it refers to identifying the type of polyp being treated [8]. However, when focusing on the application of AI techniques, classification is associated with gastrointestinal diseases that may be identified, including polyps as one of the possible conditions of the intestinal tract [9].

3.3 Research Question 1. What Artificial Intelligence Techniques Have Been Proposed for the Identification and Segmentation of Intestinal Polyps?

The techniques reviewed in the 26 analyzed papers follow a common pattern, which is structured into four main components: (1) the architecture, which includes a backbone accompanied by additional processing; (2) preprocessing of the image dataset; (3) training using one or more datasets; and (4) evaluation using performance metrics.

Since there is no standard model and the characteristics of each proposal are so varied, it is only possible to examine the backbone in depth to relate the different methods to one another. Therefore, this research question, in order to be answered, focuses exclusively on the backbone component, as it is the only common characteristic that allows for comparison and evaluation across the methods proposed.

The backbone of the architectures constitutes the process through which features are extracted from the data, with pre-trained Artificial Neural Networks (ANNs) being a common prototype. In this review, three types of models were identified: CNN (Convolutional Neural Network), a type of neural network that uses convolution layers to detect local patterns (such as edges, textures, or shapes) [10]; Transformer, a type of neural network designed to handle data sequences, such as text or time series. It uses attention mechanisms for both global and local approaches [11]; and CNN + Transformer, the integration of both backbones. Figure 2(a) shows the recurrence of each of these backbones in the reviewed papers.

Another notable feature regarding the backbones is the frequency with which each one appears over the years. The most recurrent backbone is CNN, as it is the oldest; however, over time, it has been shown that the use of other architectures provides better performance [12], such as Transformer or the combination of CNN with Transformer, as seen in Figure 2(b). It is worth noting that no papers from 2020 were found in the selection for this review.

In addition to the base architecture, preprocessing is also performed in each of the techniques to process the images from the datasets and unify their characteristics. These modifications, shown in Figure 3, are mentioned in 17 out of the 26 papers, and the following were identified: data augmentation, image normalization, recoloring, pixel resizing, and CLAHE (contrast-limited adaptive histogram equalization). Moreover, in 9 papers, preprocessing is mentioned but not described, as shown in Figure 3 as “Not specified”.

The data augmentation is worth mentioning, defined as a technique that involves rotation, scaling, horizontal and vertical flipping, and translation of each image to generate new samples and increase the dataset used for model training [9], in order to mitigate the overfitting problem; this occurs when the model performs exceptionally well on the training set but shows poor performance on the test set or unseen data [13].

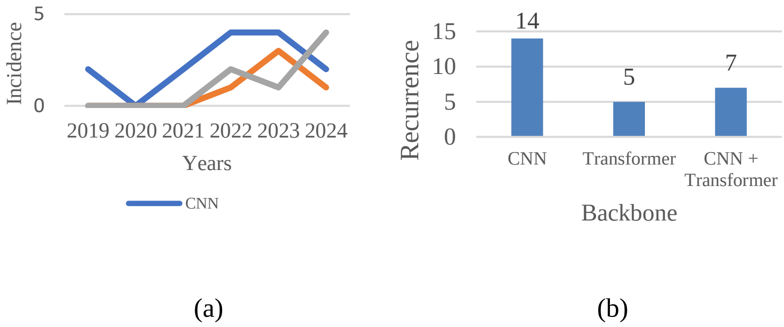


Fig. 2. (a) Recurrence of the identified backbones. (b) Incidence of the use of backbones over the years 2019-2024.

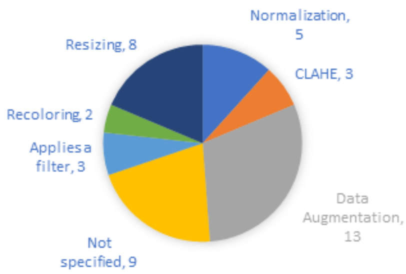


Fig. 3. Modifications made in the preprocessing of the dataset's images.

The datasets, although they will be analyzed in detail in section 3.4 of the review, have very varied characteristics, so each technique must apply this treatment to the images before passing them through the proposed model for training. An important relationship was found between image dimensions and model performance [14]. In Figure 4, we can observe the most common dimensions (pixels) and their frequency of use, as specified in 18 of the 26 selected papers.

3.4 Research Question 2. What are the Characteristics of the Image Datasets Available in the Literature for the Classification of Intestinal Polyps?

The image and video datasets used in the reviewed studies come from colonoscopies. In addition to the polyp image, these datasets include manual annotations made by specialists, who marked the location of the polyp with a bounding box or a segmentation mask, both called "Ground Truth" [10].

For the analysis, only those datasets that were used at least twice within the reviewed papers were considered. These include CVC-ClinicDB, CVC-ColonDB, ETIS-LaribPolypDB, Kvasir-SEG, Kvasir, Kvasir-Sessile, and Hyper Kvasir, with its frequency of appearance shown in Figure 5.

The most frequently used dataset was CVC-ClinicDB, followed by CVC-ColonDB and ETIS-LaribPolypDB. As shown in Table 2, all relevant datasets were created between 2012 and 2021 and are in JPG format, except for CVC-ClinicDB, which is in PNG/TIF format. Most datasets use the WLI (White Light Imaging) modality,

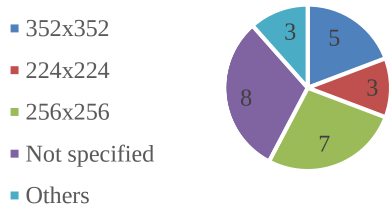


Fig. 4. Frequency of use of the selected dimension in image preprocessing**.**

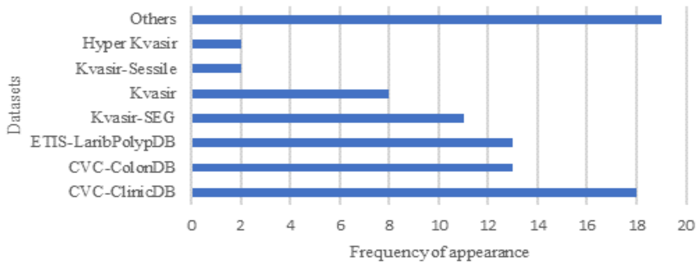


Fig. 5. Frequency of appearance of the most recurrent datasets.

including Kvasir and HyperKvasir, which also incorporate NBI (Narrow Band Imaging). However, CVC-ColonDB and Kvasir-Sessile do not specify the imaging modality.

WLI refers to capturing images using white light, a standard technique in colonoscopy [15]. In contrast, NBI is an advanced imaging technique that enhances the visualization of mucosal and vascular patterns by using narrow-band filters, improving the detection and characterization of lesions [11].

One particularity that can be highlighted about these datasets is the number of images they contain, as it is considered quite limited for the model training stage. This may explain why, in the image preprocessing, data augmentation is applied in 13 of the 26 papers analyzed.

In some studies, such as Pham et al. [16] with “seUNet-Trans” or Saad et al. [17] with “PolySeg Plus”, propose combining multiple datasets to increase the amount of data for both training and evaluation of the model. Additionally, an observable trend in the reviewed papers is the use of the “hold-out” validation method, which involves splitting the dataset into two parts: 80% for model training and 20% for evaluation. This allows for optimized learning and balanced performance measurement.

Since the image dimensions in the dataset do not match those used by the models, this explains the resizing process carried out in the studies during the preprocessing stage to unify their characteristics.

Table 2. Comparison of the most frequent datasets and their characteristics**.**

Name of the dataset	Year	# Images	Modality	Image size	Videos	Format
CVC-ClinicDB	2015	612	WLI	384x288	31	PNG / TIF
CVC-ColonDB	2012	380	-	574x500	15	JPG
ETIS-LaribPolypDB	2014	196	WLI	1226x996	34	JPG
Kvasir-SEG	2020	1000	WLI	332x487 - 1920x1072	-	JPG
Kvasir	2017	8000	WLI/NBI	720x576 - 1920x1072	-	JPG
Kvasir-Sessile	2021	196	-	-	-	JPG
Hyper Kvasir	2020	110,079	WLI/NBI	720x576 - 1920x1072	373	JPG

3.5 Research Question 3. What are the Most Commonly Used Performance Metrics to Evaluate Popular Artificial Intelligence Techniques in the Identification and Segmentation of Polyps?

To answer this research question, it is important to reiterate that the term “classification” refers to a different type of AI technique. While these models also identify polyps, their performance metrics are not directly comparable, as they evaluate the identification of various gastrointestinal diseases rather than solely polyp detection. For this reason, papers focused on classification will not be considered in this section, as they do not present enough common metrics to conduct a proper performance analysis. Consequently, the study will focus exclusively on the remaining 20 papers.

The identified metrics (with their frequency of occurrence in parentheses) from the 20 considered papers were: mDSC (15), mIoU (14), precision (14), recall (12), accuracy (5), F Score (5), F1 Score (2), F2 Score (1), MAE (3), S measure (2), specificity (2), and sensitivity (1). Among these, only the first five were considered to describe, analyze, relate, and evaluate the proposed models.

Each of these metrics evaluates different aspects of the model, making them unsuitable for direct comparison. Instead, they are selected based on the specific needs of each proposed technique.

The metric (a) mDSC (mean Dice Similarity Coefficient) refers to the relationship between the Ground Truth and the model’s prediction overlap [10]; (b) mIoU (mean Intersection over Union) represents the average ratio between the predicted area and the actual area [10]; (c) precision measures the proportion of correctly predicted positive cases out of the total positive predictions made by the model [18]; (d) accuracy is the proportion of correct predictions, including both true positives and true negatives, over the total predictions made by the model [19]; and finally, (e) recall describes the proportion of actual positives that were correctly identified [20].

Below are the respective formulas for these performance metrics, where TP (True Positives) are correctly detected polyps, FP (False Positives) are regions mistakenly identified as polyps, TN (True Negatives) are correctly identified non-polyp areas, and FN (False Negatives) are polyps that the model failed to detect:

$$\text{mDSC} = \frac{2TP}{2TP+FP+FN}, \tag{a}$$

$$\text{mIoU} = \frac{TP}{TP+FP+FN}, \tag{b}$$

$$\text{Precision} = \frac{TP}{TP+FP}, \tag{c}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}, \tag{d}$$

$$\text{Recall} = \frac{TP}{TP+FN}. \tag{e}$$

Due to the lack of standardization in model performance evaluation metrics, the obtained assessments for each proposed technique cannot be directly compared. A link is provided in Section 4 for a comprehensive review of each technique and its corresponding performance.

4 Discussion

Among all the proposals in the 26 reviewed papers, a strong connection was identified between the chosen backbone, its evolution over the years, and its performance. The trend shows a preference for Transformer-based models or the combination of CNN with Transformer due to their positive impact on evaluation metrics. For a better understanding of the techniques and the collected information, refer to the annex.

The datasets used in the reviewed studies share the characteristic of being small, as they consist of a limited number of videos and images. Most approaches opted to preprocess the datasets to homogenize their characteristics. The "hold-out" model was used for evaluation; however, given the small size of the datasets, the use of "cross-validation" is recommended, as it provides a more robust performance estimation.

Despite these limitations, the application of AI in endoscopy has gained interest from both academia and industry, leading to the development of commercial solutions that integrate computer-aided diagnosis into clinical practice. Some companies, such as Olympus with EndoBRAIN and EndoAid [21], as well as available systems like EndoMind [22] and CADEYE by Fujifilm [21], have developed and commercialized software for this purpose, though at a high cost. However, EndoMind is open-source software, making it an accessible alternative for researchers and developers. Additionally, various competitions focused on polyp detection [23] have been promoted, enabling the development of multiple datasets and the benchmarking of different methodological approaches.

Moreover, the metrics used to evaluate the proposed AI techniques vary widely, assessing different aspects in each approach, leading to a lack of standardization in truly determining each model's performance. As observed in the recurrence of each metric, not even the most popular ones (mDSC, mIoU, precision, accuracy, and recall) were reported in all the reviewed papers. Additionally, the high performance observed may be influenced by the limited number of images in the datasets; while this does not constitute overfitting per se, the models might exhibit lower performance in real-world scenarios due to the insufficient number of training samples.

5 Conclusion

After reviewing 26 selected papers using the Kitchenham and Charters methodology, the proposed objectives were achieved, and the research questions were answered.

The proposed techniques, the architecture of each, their backbone, the corresponding preprocessing, the datasets used for training, and the performance metrics used to evaluate these techniques were identified.

From the findings, it is observed that significant challenges remain, particularly the lack of standardization in performance metrics for identification and segmentation, which makes it difficult to make objective comparisons between models performing these tasks. It is recommended to use widely accepted metrics, such as sensitivity and specificity, to complement traditional metrics and improve the validity of comparisons.

Another relevant aspect is the small size of the datasets used, which may influence the high performance reported in the studies. A possible solution for future research would be the creation of larger and more diverse datasets, in addition to implementing advanced data augmentation and transfer learning techniques to improve the robustness of the models.

As future work, considering that some companies already market AI-based software for polyp detection (such as Olympus), a comparative evaluation between these commercial systems and the academic models reviewed is recommended.

References

1. Al-Shamsi, H.O., Iqbal, F., Kourie, H.R.: Colorectal Cancer in the UAE. In: Al-Shamsi, H.O. (eds) Cancer Care in the United Arab Emirates, (2024) doi: 10.1007/978-981-99-6794-0_27.
2. Duc, N.T., Oanh, N.T., Thuy, N.T.: ColonFormer: An Efficient Transformer-based Method for Colon Polyp Segmentation. In: IEEE Access, 10, pp. 80575–80586 (2022) doi: 10.1109/ACCESS.2022.3195241.
3. Peng, Y., Feng, M., Zhai, Z.: PDLFBR-Net: Partial Decoder Localization and Foreground-Background Refinement Network for Polyp Segmentation. In: IEEE Access, 12, pp. 114280–114294 (2024) doi: 10.1109/ACCESS.2024.3445428.
4. Lima, A.C. d. M., Paiva, L.F. d., Bráz, G.: A Two-Stage Method for Polyp Detection in Colonoscopy Images Based on Saliency Object Extraction and Transformers. In: IEEE Access, 11, pp. 76108–76119 (2023) doi: 10.1109/ACCESS.2023.3297097.
5. Kitchenham, B., Charters, S.: Guidelines for Performing Systematic Literature Reviews in Software Engineering Version 2.3. EBSE Technical Report EBSE-2007-01. https://www.researchgate.net/publication/302924724 (2007)
6. Guo, Q., Fang, X., Wang, L.: Polyp Segmentation of Colonoscopy Images by Exploring the Uncertain Areas. In: IEEE Access, 10, pp. 52971–52981 (2022) doi: 10.1109/ACCESS.2022.3175858.
7. Lalinia, M., Sahafi, A.: Colorectal Polyp Detection in Colonoscopy Images Using YOLO-V8 Network. Signal, Image and Video Processing, 18, pp. 2047–2058 (2024) doi: 10.1007/s11760-023-02835-1.

8. Ahamed, M.F., Islam, M.R., Nahiduzzaman, M.: Automated Detection of Colorectal Polyp Utilizing Deep Learning Methods with Explainable AI. In: IEEE Access, 12, pp. 78074–78100 (2024) doi: 10.1109/ACCESS.2024.3402818.
9. Yogapriya, J., Chandran, V., Sumithra, M.G.: Gastrointestinal Tract Disease Classification from Wireless Endoscopy Images Using Pretrained Deep Learning Model. International Journal of Biomedical Imaging, (2021) doi: 10.1155/2021/5940433.
10. Yuan, J., Liu, G., Nam, H.: Polyp Segmentation Based on Multilevel Information Correction Transformer. In: IEEE Access, 12, pp. 91619–91633 (2024) doi: 10. 1109/ACCESS.2024.3421296.
11. Sikkandar, M.Y., Sundaram, S.G., Alassaf, A.: Utilizing Adaptive Deformable Convolution and Position Embedding for Colon Polyp Segmentation with a Visual Transformer. Scientific Reports, 14, Article 7318 (2024) doi: 10.1038/s41598-024-57993-0.
12. Pham, T. -H., Li, X., Nguyen, K. -D.: seUNet-Trans: A Simple Yet Effective UNet-Transformer Model for Medical Image Segmentation. In: IEEE Access, 12, pp. 122139–122154, doi: 10.1109/ACCESS.2024.3451304.
13. Obayya, M., Al-Wesabi, F.N., Maashi, M.: Modified Salp Swarm Algorithm with Deep Learning Based Gastrointestinal Tract Disease Classification on Endoscopic Images. In: IEEE Access, 11, pp. 25959–25967 (2023) doi: 10.1109/ACCESS.2023.3256084.
14. Elkarazle, K., Raman, V., Then, P.: Improved Colorectal Polyp Segmentation Using Enhanced MA-NET and Modified Mix-ViT Transformer. In: IEEE Access, 11, pp. 69295–69309 (2023) doi: 10.1109/ACCESS.2023.3291783.
15. Zhang, M., Sun, Q., Cai, F.: MHA-Net: A Multibranch Hybrid Attention Network for Medical Image Segmentation. Computational and Mathematical Methods in Medicine, (2022) doi: 10.1155/2022/8375981.
16. Hong, L.T.T., Thanh, N.C., Long, T.Q.: CRF-EfficientUNet: An Improved UNet Framework for Polyp Segmentation in Colonoscopy Images with Combined Asymmetric Loss Function and CRF-RNN Layer. In: IEEE Access, 9, pp. 156987–157001 (2021) doi: 10.1109/ACCESS.2021.3129480.
17. Saad, A.I., Maghraby, F.A., Badawy, O.: PolySeg Plus: Polyp Segmentation Using Deep Learning with Cost-Effective Active Learning. Journal of Ambient Intelligence and Humanized Computing, 16, Article 148, (2023) doi: 10.1007/s44196-023-00330-6.
18. Kader, R., Cid-Mejias, A., Brandao, P.: Polyp Characterization Using Deep Learning and a Publicly Accessible Polyp Video Database. Dentomaxillofacial Radiology, (2022) doi: 10.1111/den.14500.
19. Srinivasan, S., Durairaju, K., Deeba, K.: Multimodal Biomedical Image Segmentation Using Multi-Dimensional U-Convolutional Neural Network. BMC Medical Imaging, 24, Article 38, (2024) doi: 10.1186/s12880-024-01197-5.
20. Kang, J., Gwak, J.: Ensemble of Instance Segmentation Models for Polyp Segmentation in Colonoscopy Images. In: IEEE Access, 7, pp. 26440–26447 (2019) doi: 10.1109/ ACCESS.2019.2900672.
21. Kamitani, Y., Nonaka, K., Isomoto, H.: Current Status and Future Perspectives of Artificial Intelligence in Colonoscopy. Journal of Clinical Medicine, 11(10), 2923 (2022) doi: 10.3390/jcm11102923.
22. Lux, T.J., Banck, M., Saßmannshausen, Z.: Pilot Study of a New Freely Available Computer-aided Polyp Detection System in Clinical Practice. International Journal of Colorectal Disease, 37(6), pp. 1349–1354 (2022) doi: 10.1007/s00384-022-04178-8.
23. Ali, S., Ghatwary, N., Jha, D.: Assessing Generalisability of Deep Learning-Based Polyp Detection and Segmentation Methods Through a Computer Vision Challenge. Scientific Reports, 14(1), Artículo 2032, (2024) doi: 10.1038/s41598-024-52063-x.
24. Al-Mekhlafi, Z.G., Senan, E.M., Alshudukhi, J.S.: Hybrid Techniques for Diagnosing Endoscopy Images for Early Detection of Gastrointestinal Disease Based on Fusion Features. International Journal of Intelligent Systems, pp. 1–20 (2023) doi: 10.1155/2023/8616939.

25. Bhattacharya, D., Eggert, D., Betz, C.: Squeeze and Multi-Context Attention for Polyp Segmentation. International Journal of Imaging Systems and Technology, 33 (2022) doi: 10.1002/ima.22795.
26. Chen, L., Ge, H., Li, J.: CrossFormer: Multi-scale Cross-Attention for Polyp Segmentation. IET Image Processing, 17 (2023) doi: 10.1049/ipr2.12875.
27. Nguyen, N. -Q., Lee, S.-W.: Robust Boundary Segmentation in Medical Images Using a Consecutive Deep Encoder-Decoder Network. In: IEEE Access, 7, pp. 33795–33808 (2019) doi: 10.1109/ACCESS.2019.2904094.
28. Raju, A.S.N., Jayavel, K., Rajalakshmi, T.: ColoRectalCADx: Expeditious Recognition of Colorectal Cancer with Integrated Convolutional Neural Networks and Visual Explanations Using Mixed Datasets Evidence. Computational and Mathematical Methods in Medicine, (2022) doi: 10.1155/2022/8723957.
29. Yasmin, F., Hassan, M.M., Hasan, M.: GastroNet: Gastrointestinal Polyp and Abnormal Feature Detection and Classification with Deep Learning Approach. In: IEEE Access, 11, pp. 97605–97624 (2023) doi: 10.1109/ACCESS.2023.3312729.
30. Younas, F., Usman, M., Yan, W.Q.: A Deep Ensemble Learning Method for Colorectal Polyp Classification with Optimized Network Parameters. Pattern Analysis and Applications, 53, pp. 2410–2433 (2023) doi: 10.1007/s10489-022-03689-9.

Annex

			[redacted]	Architecture		Image processing		Evaluation				
Title	**date**	**Academic databases**	**Proposed AI technique**	**Backbone**	**Additional processing**	**Resizing**	**Extra modifications**	**mDSC**	**mIoU**	**Precision**	**Accuracy**	**Recall**
seUNet-Trans: A Simple Yet Effective UNet-Transformer Model for Medical Image Segmentation	sep-24	IEEEXplore	seUNet-Trans	CNN + Transformer	seU-Net + Head Transformer	128x128	Dataset Combination	0.919	0.85	0.926	NS	0.912
PDLFBR-Net: Partial Decoder Localization and Foreground-Background Refinement Network for Polyp Segmentation	ago-24	IEEEXplore	PDLFBR-Net	Transformer	Cross-level Attention-enhanced Fusion Module (CAFM) + Position Recognition Module (PRM) + Foreground-Background Refinement Module (FBRM)	352x352	Data augmentation	0.937	0.895	NS	NS	NS
Polyp Segmentation Based on Multilevel Information Correction Transformer	jul-24	IEEEXplore	MICT	CNN + Transformer	Multiscale Feature Extractor (MFE) + Multilevel Lesion Correction Module (MLC) + Feature Selection Fusion Module (FSF)	352x352	NS	0.8176	0.7422	NS	NS	NS
Automated Detection of Colorectal Polyp Utilizing Deep Learning Methods With Explainable AI	may-24	IEEEXplore	TR-SE-Net	CNN + Transformer	ResNet50 + Squeeze-and-Excitation Blocks	256x256 y 224x224	Contrast Limited Adaptive Histogram Equalization (CLAHE) + data augmentation	0.875	0.7961	0.9027	NS	0.8879
Automated Colorectal Polyps Detection from Endoscopic Images using MultiResUNet Framework with Attention Guided Segmentation	abr-24	SpringerLink	MultiResU-Net Framework	CNN	MultiResUNet Blocks + Guided Attention (GA)	256x256	Normalization + Data Augmentation	0.8663	0.8277	0.9364	0.9593	0.806
Utilizing adaptive deformable convolution and position embedding for colon polyp segmentation with a visual transformer	mar-24	SpringerLink	Polyp-ViT	CNN + Transformer	ViT + Adaptive Deformable Convolutional Network (ADCN) + ResNet Blocks	256x256	Images in JPG format	0.9871	0.9889	0.9827	0.9891	NS
Multimodal Biomedical Image Segmentation using Multi-Dimensional U-Convolutional Neural Network	feb-24	SpringerLink	MDU-CNN	CNN	U-Net + Multidimensional Convolutions + Skip Connections + Convolution Paths	256x256	NS	NS	NS	NS	NS	NS

			[redacted]	Architecture		Image processing		Evaluation				
Colorectal polyp detection in colonoscopy images using YOLO-V8 network	dic-23	SpringerLink	YOLO-V8 for polyp detection	CNN + Transformer	YOLOv8 + bounding boxes + Extremely Lightweight Adaptive Networks + Task-aligned One-stage Object Detection	NS	Normalization + Filters (Hue and Brightness) + Data Augmentation	NS	NS	0.956	NS	0.917
GastroNet: Gastrointestinal Polyp and Abnormal Feature Detection and Classification With Deep Learning Approach	sep-23	IEEEXplore	GastroNet	CNN	YOLOv5 + Cross Stage Partial Networks (CSPDarknet) + Neck: PANet (Path Aggregation Network) y SPPF (Spatial Pyramid Pooling - Fast)	416x416	Normalization + Data (Mosaic) Augmentation	NS	NS	0.99 * (NA)	NS	1 * (NA)
PolySeg Plus: Polyp Segmentation Using Deep Learning with Cost Effective Active Learning	ago-23	SpringerLink	PolySeg Plus	CNN	Unet/UNet++/ResUNet++ /ResUNet + Locally Shared Features (LSF) + grid search	224x224	Data Augmentation + Gaussian Filters	0.9476	NS	0.8768	NS	0.9245
CrossFormer: Multi-scale cross-attention for polyp segmentation	jul-23	Wiley	CrossFormer	Transformer	PVTv2 (Pyramid Vision Transformer v2) + CEM (Channel Enhancement Module) + MSCAM (Multi-Scale Cross-Attention Module)	352x352	NS	0.9249	0.8739	0.9259	NS	0.9437
Improved Colorectal Polyp Segmentation Using Enhanced MA-NET and Modified Mix-ViT Transformer	jul-23	IEEEXplore	Enhanced MA-NET	Transformer	Multi-Scale Attention Network (MA-NET) + Mix-ViT	256x256	Normalization + CLAHE + CIELAB Color Conversion	0.983	0.973	0.989	NS	0.983
A Two-Stage Method for Polyp Detection in Colonoscopy Images Based on Saliency Object Extraction and Transformers	jul-23	IEEEXplore	SOE DETR	Transformer	Detection Transformer (DETR) + Dense Prediction Transformer (DPT) + Visual Saliency Transformer (VST)	640x640	Images converted to SGB (Grayscale)	NS	NS	0.932	0.842	0.879
Hybrid Techniques for Diagnosing Endoscopy Images for Early Detection of Gastrointestinal Disease Based on Fusion Features	abr-23	Wiley	VGG-16 / DenseNet-121	CNN	VGG-16/DenseNet-121 + SVM (Support Vector Machines)/Random Forest + PCA	NS	Data augmentation	NS	NS	NS	NS	NS
Modified Salp Swarm Algorithm With Deep Learning Based Gastrointestinal Tract Disease Classification on Endoscopic Images	mar-23	IEEEXplore	MSSADL-GITDC	CNN	CapsNet (Capsule Network) + Class Attention Layer (CAL) + Modified Salp Swarm Algorithm (MSSA) + DBN (Deep Belief Network) + ELM (Extreme Learning Machine)	NS	Median Filter (MF)	NS	NS	0.9216 (NA)	0.9803 (NA)	0.9213 (NA)
Polyp characterization using deep learning and a publicly accessible polyp video database	dic-22	Wiley	CNN Classification	CNN	ResNet101 + Softmax	768x576	NS	NS	NS	NS	NS	NS
ColoRectalCADx: Expeditious Recognition of Colorectal Cancer with Integrated Convolutional Neural Networks and Visual Explanations Using Mixed Dataset Evidence	nov-22	Wiley	ColoRectalCADx	CNN	ResNet-50/VGG-16/DenseNet-201 + Support Vector Machine (SVM)/Long Short-Term Memory (LSTM) + Grad-CAM	224x224	Data augmentation	NS	NS	NS	NS	NS
MHA-Net: A Multibranch Hybrid Attention Network for Medical Image Segmentation	oct-22	Wiley	MHA-Net	CNN + Transformer	ResNet-34 + MA-NET + Mix-ViT (Mix Transformer) + Multiscale Attention Module (PSA)	256x256	Normalization + CLAHE	0.7692	0.8311	0.8618	0.9656	0.8
ColonFormer: An Efficient Transformer Based Method for Colon Polyp Segmentation	ago-22	IEEEXplore	ColonFormer	CNN + Transformer	Mix Transformer (MiT) + UPerNet + Reverse Attention	352x352	NS	0.927	0.877	NS	NS	NS
Squeeze and multi-context attention for polyp segmentation	jul-22	Wiley	Squeeze and Multi-Context Attention	CNN	U-Net/Attention U-Net/R2U-Net/ResUNet++/R2AU-Net + SMCA (Squeeze Multi-Context Attention)	196x196, 256x256 (test) y 512x512 (training)	NS	0.58	0.47	0.57	NS	0.78
Polyp Segmentation of Colonoscopy Images by Exploring the Uncertain Areas	may-22	IEEEXplore	UnX + FeE	Transformer	Uncertainty eXploration (UnX) + Feature Enhancement (FeE)	352x352	NS	0.912	0.859	NE	NS	NS
A deep ensemble learning method for colorectal polyp classification with optimized network parameters	may-22	SpringerLink	Deep ensemble learning method	CNN	GoogLeNet/ResNet-50/Inception-v3/Xception/DenseNet-201/SqueezeNet + transfer learning	NS	Data Augmentation + Data Oversampling	NS	NS	NS	NS	NS
CRF-EfficientUNet: An Improved UNet Framework for Polyp Segmentation in Colon-	nov-21	IEEEXplore	CRF-EfficientU-Net	CNN	Unet + CRF-RNN (Conditional Random Field as a Recurrent Neural Network) + EfficientNet B7	NS	Data Augmentation	0.9272	0.8769	0.9492	NE	0.9702

			[redacted]	Architecture		Image processing		Evaluation				
oscopy Images With Combined Asymmetric Loss Function and CRF-RNN Layer												
Gastrointestinal Tract Disease Classification from Wireless Endoscopy Images Using Pretrained Deep Learning Model	sep-21	Wiley	Wireless Endoscopy	CNN	VGG16/ResNet-18/GoogleNet	NS	Data Augmentation	NS	NS	NS	NS	NS
Robust Boundary Segmentation in Medical Images Using a Consecutive Deep Encoder-Decoder Network	mar-19	IEEEXplore	CDED-net	CNN	Deep Encoder-Decoder Networks (DEDNs) + Dicoss-loss	NS	Data Augmentation	0.891	NS	0.95	0.987	NS
Ensemble of Instance Segmentation Models for Polyp Segmentation in Colonoscopy Images	feb-19	IEEEXplore	Ensemble Mask R-CNN	CNN	ResNet-50/ResNet-101 + Two R-CNN Masks	NS	Data Augmentation	NS	0.6946	0.7792	NS	0.7625
								Average of the obtained evaluation				
								0.881684615	0.808990909	0.89555	0.94435	0.877618182

Notes:

- Marked in blue: Refers to the articles specialized in gastrointestinal disease classification.
- Marked in yellow: highest scores achieved on each evaluation metric.
- NS: Refers to “Not Specified”.
- NA: Refers to “Not Applicable”.