Intelligence-Based Medicine 12 (2025) 100262

Contents lists available at ScienceDirect

Intelligence-Based Medicine

journal homepage: www.sciencedirect.com/journal/intelligence-based-medicine

Advancements in artificial intelligence for colorectal cancer: A comprehensive overview of systematic reviews

Aurea Valeria Pereira Silva *, Plinio Sa Leitao-Junior

Instituto de Informática - Universidade Federal de Goiás, Alameda Palmeiras, s/n - Chácaras Califórnia, Goiânia, 74690-900, Goiás, Brazil

ARTICLE INFO

Keywords:
Colorectal cancer
Systematic review
Computational intelligence
Diagnosis and treatment of colorectal cancer

ABSTRACT

Background: Colorectal cancer (CRC) is a leading cause of cancer-related mortality worldwide. Computational intelligence (CI) has emerged as a promising tool to improve diagnosis, staging, and treatment, but evidence remains scattered across the literature.
Objective: This tertiary review aims to synthesize systematic reviews on CI applications in CRC care, highlighting algorithms, datasets, performance metrics, clinical scopes, and methodological gaps.
Methods: A structured search in PubMed and EMBASE identified systematic reviews published between 2018 and 2023, following PRISMA guidelines. Twenty-two reviews were included. Extracted data covered CI techniques, evaluation methods, target outcomes, and dataset characteristics. Risk of bias was assessed using AMSTAR 2, and overlap of primary studies was analyzed through a correlation matrix.
Results: The reviews addressed four clinical scopes: macroscopic lesion classification (colonoscopy), histological analysis, disease staging, and survival or treatment prediction. Convolutional neural networks (CNNs) were the most commonly used models. While some applications showed high performance (AUC > 0.90), most reviews had low to moderate methodological quality. Key limitations included lack of external validation, dataset heterogeneity, and limited generalizability. Significant overlap was observed in studies focused on colonoscopy-based tasks.
Conclusion: CI offers valuable contributions to CRC management, but broader clinical adoption is hindered by methodological inconsistencies and insufficient validation. This review provides a comprehensive synthesis to guide future research and promote the development of robust, explainable, and generalizable models for clinical use.

1. Introduction

Colorectal cancer (CRC) is one of the most common types of cancer, with symptoms such as changes in bowel habits, rectal bleeding, and abdominal pain. Risk factors include family history, advanced age, and unhealthy lifestyle habits [1]. According to the World Health Organization (WHO), it is the third most prevalent cancer globally, with approximately 1.93 million new cases reported in 2020 [2]. Early detection is crucial for achieving better clinical outcomes, and the main screening methods include colonoscopy, sigmoidoscopy, and fecal occult blood tests [3].

With the advancement of computational intelligence (CI), new possibilities have emerged to support the diagnosis, staging, and treatment of colorectal cancer. Studies have shown that machine learning algorithms are effective in tasks such as polyp detection, histological analysis, survival prediction, and clinical decision support [4]. However, challenges related to model validity and generalizability still limit their widespread application. Given the growing scientific output on the topic, it is essential to critically organize this body of knowledge by identifying patterns, advancements, methodological gaps, and opportunities for clinical application. In this context, literature reviews play a fundamental role in synthesizing the findings of multiple primary studies.

Tertiary reviews, also known as overviews of reviews, are a type of study that systematically collects and evaluates the findings of previously published secondary reviews. Their purpose is to provide a broad and integrated perspective on a research field, identifying convergences, divergences, methodological patterns, and gaps that still exist in the literature.

In this context, it is essential to understand how the scientific literature has addressed these issues. This tertiary review addresses a previously unexplored gap by investigating systematic reviews on the use of machine learning in colorectal cancer—an area not yet examined by other tertiary studies. This originality reinforces the relevance of the

* Corresponding author.
E-mail addresses: aureavaleria8@gmail.com (A. Valeria Pereira Silva), plinio@inf.ufg.br (P. Sa Leitao-Junior).

https://doi.org/10.1016/j.ibmed.2025.100262
Received 11 July 2024; Received in revised form 17 May 2025; Accepted 27 May 2025
Available online 19 June 2025

study, which aims to organize and synthesize evidence in an emerging field with significant public health impact.

Unlike a traditional systematic review, this study adopts a broader analytical scope by mapping methodological trends, applied algorithms, evaluation metrics, and datasets used across existing reviews. In addition, the methodological risk of bias was assessed using the AMSTAR 2 tool, and the relationships between studies were visualized through graphs, enabling a deeper understanding of the field's evolution and research gaps.

Section 1 presents the clinical context of colorectal cancer, the advancements in computational intelligence, and the rationale for conducting this tertiary review. Section 2 details the methodology employed, including selection criteria, data extraction, bias assessment, and use of a correlation matrix. Section 3 describes the screening results and characterizes the included studies. Section 4 organizes the main findings, focusing on algorithms, evaluation metrics, and datasets. Finally, Section 5 critically discusses the evidence gathered, highlighting clinical implications, methodological limitations, and future directions for the application of CI in colorectal cancer.

2. Methods

This section outlines the adherence to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines in conducting the systematic review. PRISMA consists of 27 checklist items crucial for maintaining research quality and ensuring the acceptance of results in scientific publications [5]. Emphasizing transparency and reliability, PRISMA's structured approach in this review enhances methodological validity and result credibility.

2.1. Literature search strategy

A review was conducted to identify literature reviews on AI algorithms in colorectal cancer diagnosis and treatment, assessing their effectiveness and risks. Searches were limited to article titles in PUBMED (MEDLINE) and EMBASE databases from January 1, 2018, to December 31, 2023, using specific search terms for pathology, interventions, and study types.

- Scope: ("adenocarcinoma" OR "cancer" OR "carcinoma" OR "carcinomas" OR "sarcoma" OR "sarcomas" OR "metastasis" OR "neoplasia" OR "tumor" OR "tumors" OR "tumour" OR "tumours") AND ("colon" OR "rectal" OR "colorectal")
- Intervention: AND ("bayes" OR "cnn" OR "computational intelligence" OR "convolutional network" OR "convolutional networks" OR "decision tree" OR "decision trees" OR "deep learning" OR "knn" OR "k-nearest neighbors" OR "learning analytics" OR "linear regression" OR "logistic regression" OR "machine learning" OR "nearest neighbors" OR "neural network" OR "neural networks" OR "random forest" OR "rnn" OR "svm" OR "support vector machine")
- Classification and categorization process: AND ("literature map" OR "literature mapping" OR "literature review" OR "literature survey" OR "systematic literature" OR "systematic mapping" OR "systematic review").

2.2. Selection criteria

The first author reviewed full texts after assessing abstracts for inclusion criteria. If abstracts were insufficient, full texts were obtained and discussed with the second author to finalize decisions. Only literature reviews meeting specific criteria were included: published after 2018 in English, focusing on computational intelligence in colorectal cancer care, excluding non-literature review articles and conference abstracts. These criteria ensure study quality and minimize selection bias.

2.3. Data extraction

Following PRISMA guidelines, reviewers used a spreadsheet to collect data from studies, categorizing articles by medical issues. Data on general article information (country, publication year, cancer type), protocol specifics (search period, methodology, bias assessment), and study objectives (disease stage, data types analyzed) were extracted. This critical information helps understand study focus, evaluate results' relevance, identify gaps, and guide future research. Reviews identified three main result categories: public data sources, computational intelligence algorithms, and evaluation metrics. Public data details, including access, types, sizes, acquisition models, and references, were also documented.

2.4. Correlation matrix of articles

We created an evidence matrix to improve the assessment of primary study overlap in Literature Reviews (LR). Using it, we classified the most frequent articles in LR, identifying studies and AI related to colorectal cancer. This approach mapped relevant areas and possible gaps in the literature. The methodology provides a comprehensive and reliable way to assess the quality of studies in LR, assisting future research in the field.

2.5. Risk of bias assessment

A critical evaluation of the methodological quality of the studies included in this review was performed using AMSTAR 2, a recognized instrument for assessing systematic reviews of randomized and non-randomized studies. This assessment allowed for an objective evaluation of the quality of the included studies and identification of the overall quality of the results of the selected systematic reviews. AMSTAR 2 is widely recognized as a useful tool in the assessment of the methodological quality of systematic reviews.

3. Results

We found 49 duplicate results out of 137 in our initial searches. Of the 88 articles analyzed, nine articles with a focus other than colorectal cancer were excluded, 13 literature reviews that did not search for studies using artificial intelligence, and 23 that were not literature reviews but primary studies. Of the 43 articles that were read in full, we excluded four that did not research colorectal cancer as the main focus, we also excluded nine that did not search for studies using artificial intelligence, and two that were primary studies. We found five conference abstracts and one article in a language other than English. In the end, we included 22 literature reviews that met our eligibility criteria. Fig. 1 presents the PRISMA flowchart, which provides an overview of the secondary studies selection process.

4. Findings

This section details the selection and information extraction process from articles focusing on the application of artificial intelligence in colorectal cancer. Data were categorized into four sections: general information, protocol details, context of analyzed articles, and systematic review outcomes. The information was compiled in Table 1 - Data Extraction.

Table 1
Data extraction.

Category	Item	Description
General information	Country	Country of the first author affiliation.
	Year	Year of publication of the analyzed review.
	Type of cancer	The type of cancer being investigated.
Protocol of the reviews analyzed	Time frame	Definition of the result domains and the period of investigation for which the data were found.
	Method	The chosen review makes use of some methodology such as PRISMA for its development.
	Risk assessment tool	Tool(s) (and version) used to assess the risk of bias in the included studies.
	Database	Which bibliographic databases were searched, (eg MEDLINE, PUBMED, EMBASE).
	Studies found	Number of articles included in the review, based on the analysis of the works found.
	Studies included	How many articles were included in the review based on the analysis of the works found.
Context of the reviews analyzed	Disease stage	Assessment of the degree of spread of the disease.
	Scope	The challenge presented by the literature review.
Reported outcome of the reviews analyzed	Dataset	Public datasets found.
	Algorithms	The main algorithms used in the primary studies selected by the reviews.
	Evaluation metrics	Metrics to verify the results of the proposed methods.

4.1. General information

Two graphs (Figs. 2(a) and 2(b)) summarize general findings from the systematic reviews. Fig. 2(a) illustrates the countries where research was conducted, with a majority (17 out of 22) in high-income countries, including three studies from the United States published between 2019 and 2023. Fig. 2(b) categorizes the number of articles by the type of cancer studied. Among the 22 reviews, three focused on gastric cancer but primarily searched for primary studies on colorectal cancer. Sixteen reviews specifically included studies on AI applications in colorectal cancer, while three separately explored colon and rectal cancers.

4.2. Protocol of selected reviews

Tables 2 and 3 outline the methodologies employed in the selected literature reviews. PRISMA guidelines were predominant, utilized in 18 studies, while one review used the MOOSE checklist. Eleven studies employed bias assessment tools, with QUADAS-2 being the most frequently used. Highlighted databases included MEDLINE (PubMed), Cochrane Library, and EMBASE, with search periods spanning from 2000 to 2023, though 11 reviews had an inclusion rate below 1%.

4.3. Context of the reviews analyzed

Table 4 presents key information for understanding the applicability and generalizability of the findings from the analyzed reviews. It includes the names of the authors, publication years, disease stages, and research scopes covered by each study. These elements are essential for critically assessing the results and identifying gaps in the application of artificial intelligence (AI) in the context of colorectal cancer.

The classification of review scopes was organized into the following categories: *Stage of the Disease*, *Microscopic Features in Histology*, *Macroscopic Features at Colonoscopy*, and *Survival Prognosis and Treatment of Patients with Colorectal Cancer*. This categorization was adapted from the framework proposed by Short et al. [28], which focused on the medical cycle of diagnosis and treatment in colorectal cancer. In this study, the classification was adjusted to emphasize reviews employing computational intelligence approaches, aligning the categories with the most relevant technological strategies found in the literature.

The definitions of these categories are described below:

4.3.1. Microscopic features in histology

Histopathological analysis plays a critical role in CRC diagnosis. Four reviews focused on CI applications in histology, highlighting the use of convolutional neural networks (CNNs) for classifying tissue slides with high accuracy [7,13,20,21]. These models assist in reducing observer variability and identifying molecular features such as microsatellite instability (MSI-H) and tumor microenvironment (TME) patterns. Challenges include the lack of standardized annotations and the limited availability of large, diverse datasets, which impact reproducibility and generalizability [26].

4.3.2. Macroscopic features at colonoscopy

Macroscopic features of colonic lesions observed during colonoscopy or surgery refer to their visual appearance, which helps identify abnormalities suggestive of malignancy [28]. Nine reviews addressed polyp classification using colonoscopy images [8–12,14,19,23]. However, challenges such as real-time classification, detection of interval cancers, and limited datasets still affect the performance of AI algorithms.

Tumors may be macroscopically classified based on surface appearance into four categories: normal, reactive, nodular, and colloidal [29].

4.3.3. Stage of the disease

Staging colorectal cancer involves determining the extent of tumor progression and spread, which is critical for therapeutic planning and clinical decision-making. This process is typically performed using computed tomography (CT) and magnetic resonance imaging (MRI) [6]. Two reviews focused on this topic [6,17], discussing both traditional staging methods and the potential of machine learning techniques for automatically analyzing medical imaging data to predict cancer stage and treatment response. Staging is typically categorized into four stages: Stage I, Stage II, Stage III, and Stage IV.

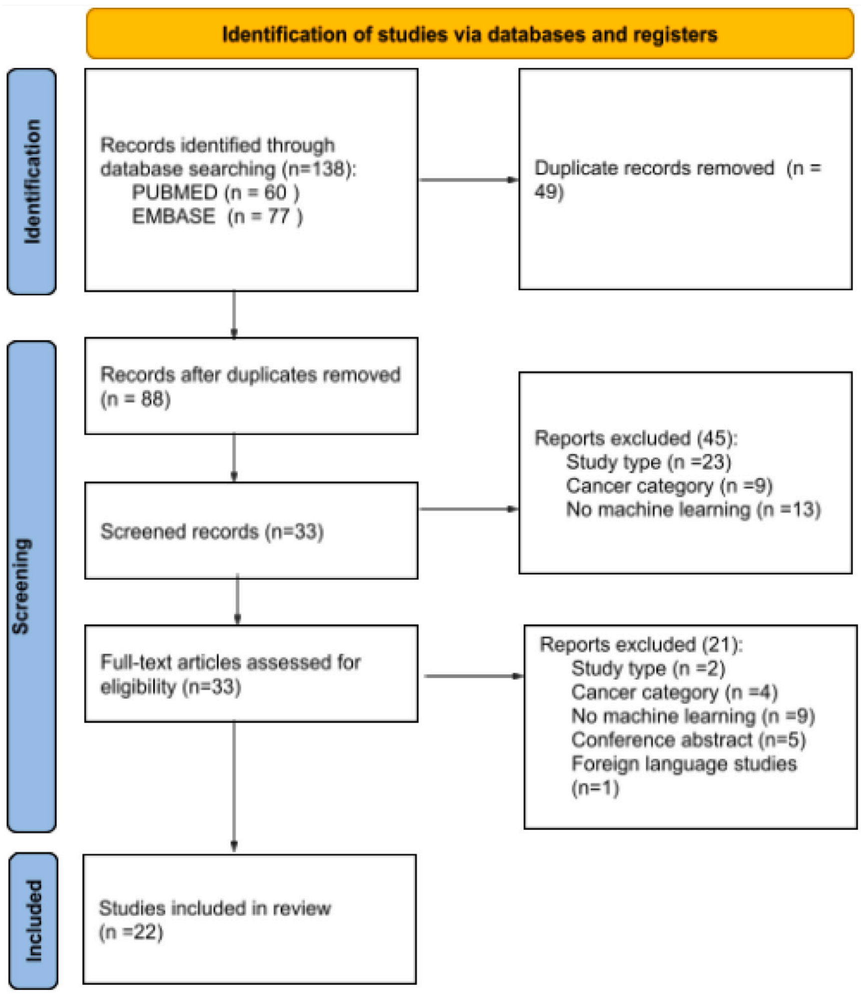
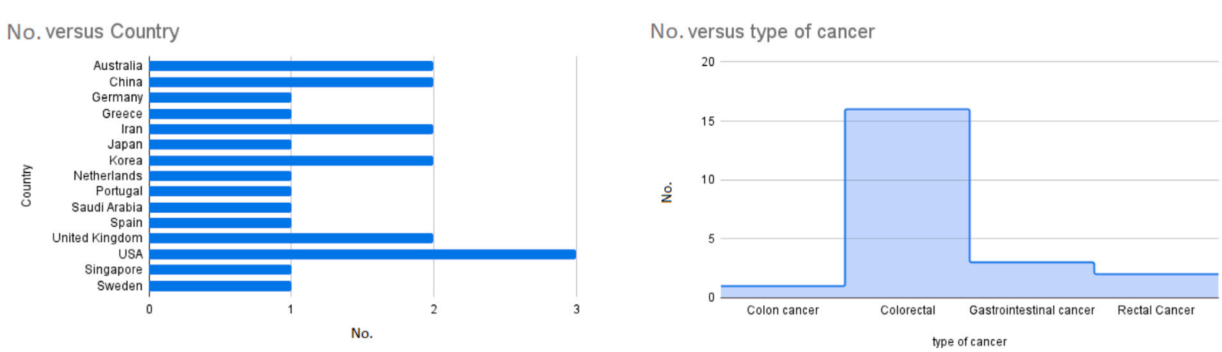


Fig. 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for study selection.



(a) Distribution of articles by countries

(b) Distribution of articles by type of cancer

Fig. 2. General information of selected literature reviews.

4.3.4. *Survival prognosis and treatment of patients with colorectal cancer*

Machine learning methods have been used to analyze complex data from medical images, histopathological slides, and electronic health records to identify patterns associated with disease progression and treatment response. These techniques support personalized care and clinical decision-making, particularly in advanced cancer stages [22, 24]. Four reviews focused on computational intelligence applied to electronic health records highlighted the need to improve data quality and develop more accurate and robust models to enhance clinical support [15,24,25,27].

4.3.5. *Cancer screening and surveillance*

Colorectal cancer screening and surveillance encompass essential processes for detecting the disease at early stages and monitoring at-risk or previously treated individuals. Screening aims to identify cancer in asymptomatic individuals using fecal occult blood tests, imaging techniques, and endoscopic exams [28]. Surveillance focuses on the

Table 2
General information extracted from selected literature reviews - part 1.

Autor, year	Search start	End of searches	Method	Risk of bias tools	Database	Studies found	Studies included
Bedrikovetski et al. [6]	January 1, 2010	October 1, 2020	PRISMA	QUADAS-2	Cochrane, PubMed/ Medline EMBASE and IEEE Xplore	65	12
Thakur et al. [7]	January 2000	January 2020	–	–	PubMed/Medline, Cochrane, and EMBASE	7503	40
Xu et al. [8]	–	April 30, 2020	PRISMA	QUADAS-2	PubMed/Medline, Web of Science, Cochrane, and EMBASE	43	13
Sanchez-Peralta et al. [9]	January 2015	February 2, 2019	PRISMA	–	ACM Digital Library, IEEE Xplore, Web of Science, PubMed, Scopus and Springer	1332	35
Nazarian et al. [10]	–	October 2020	PRISMA	QUADAS-2	EMBASE, PubMed/Medline, and Cochrane	899	48
Hassan et al. [11]	–	March 31, 2020	PRISMA	RoB 2	PubMed/Medline, EMBASE, and Cochrane	554	5
Yamada et al. [12]	January 2016	December 2021	–	–	PubMed/Medline	130	9
Kuntz et al. [13]	January 2015	December 2020	PRISMA	–	PubMed/Medline	790	16
Aziz et al. [14]	–	January 8, 2020	PRISMA	RoB 2	PubMed/Medline, EMBASE, Cochrane, Web of Science, and Computers & Applied Sciences Complete.	794	3
Liang et al. [15]	–	March 2018	PRISMA	–	PubMed, Cochrane, EMBASE and Google Scholar.	1043	7
Mohan et al. [16]	–	November 2019	MOOSE	GRADE	ClinicalTrials.gov, EBM Reviews, PubMed/Medline, EMBASE, Scopus and Web of Science.	4245	19
Moreira et al. [17]	–	–	–	–	–	–	–
Langarizadeh et al. [18]	–	December 2020	PRISMA	–	Google Scholar, Scopus, ProQuest, PubMed/Medline, Web of Science, Cochrane, and SID as a Persian database	774	44
Moen et al. [19]	–	January 4, 2022.	PRISMA	QUADAS-2	Embase, Web of Science, MEDLINE and Cochrane	1235	9
Davri et al. [20]	–	December 31, 2021	PRISMA	–	PubMed	166	82
Alam et al. [21]	–	August 2021	PRISMA	–	PubMed/Medline, EMBASE, and Cochrane	13 049	14

continuous monitoring of high-risk populations to detect recurrences or progression. Two reviews explored machine learning techniques for screening using medical images [16,18], highlighting the potential of these approaches to enhance early detection. However, machine learning still faces challenges such as the need for large datasets and difficulty in handling lesion variability, which limits its practical application.

4.4. Reported outcome of the reviews analyzed

This section goes beyond secondary studies, extracting data from primary studies cited in these reviews. It gathers diverse information, utilizing various databases and metrics to evaluate results. Hence, direct comparison for identifying the most efficient method is not suitable. The objective is to present key details from primary studies,

Table 3
General information extracted from selected literature reviews - part 2.

Autor, year	Search start	End of searches	Method	Risk of bias tools	Database	Studies found	Studies included
Alsadhan et al. [22]	January 2010	December 2020	PRISMA	JBICA tool	PubMed/Medline, EMBASE, Web of Science, and Cochrane	5348	145
Keshtkar et al. [23]	January 1, 2010	July 31, 2020	PRISMA	RoB 2	PubMed/MEDLINE, SCOPUS, Web of Science, IEEE, Inspec, ProQuest, Google Scholar, Microsoft Academic Search, ScienceOpen, arXiv, and bioRxi	3060	24
Seow-En et al. [24]	–	up to May 2023	PRISMA	PROBAST	PubMed/Medline, EMBASE, Web of Science, and Cochrane	141	7
Wang et al. [25]	January 2010	December 2020	PRISMA	QUADAS-2	PubMed/Medline, EMBASE, Web of Science, and Cochrane	97	12
Vadhwana et al. [26]	1966	6 August 2023	PRISMA	–	EMBASE and MEDLINE .	5622	6
Prabhakaran et al. [27]	–	up to July 2022	PRISMA	–	Embase (PubMed, MEDLINE) and Google Scholar	85	12

Table 4
Context extraction of selected literature reviews.

Autor, year	Disease stage	Scope
Bedrikovetski et al. [6]	Lymph node metastasis	Stage of the disease
Thakur et al. [7]	–	Microscopic features in histology
Xu et al. [8]	Polyps of any size	Macroscopic features at colonoscopy
Sanchez-Peralta et al. [9]	–	Macroscopic features at colonoscopy
Nazarian et al. [10]	–	Macroscopic features at colonoscopy
Hassan et al. [11]	–	Macroscopic features at colonoscopy
Yamada et al. [12]	Flat lesions	Macroscopic features at colonoscopy
Kuntz et al. [13]		Microscopic features in histology
Aziz et al. [14]		Macroscopic features at colonoscopy
Liang et al. [15]	–	Survival prognosis and treatment of patients with colorectal cancer
Mohan et al. [16]	–	Cancer screening and surveillance
Moreira et al. [17]	–	Stage of the disease
Langarizadeh et al. [18]	–	Cancer screening and surveillance
Moen et al. [19]	–	General
Davri et al. [20]	–	Microscopic features in histology
Alam et al. [21]	Solid Cancers	Microscopic features in histology
Alsadhan et al. [22]	–	Survival prognosis and treatment of patients with colorectal cancer
Keshtkar et al. [23]	Polyps of any size	Macroscopic features at colonoscopy
Seow-En et al. [24]	Locally advanced cancer	Survival prognosis and treatment of patients with colorectal cancer
Wang et al. [25]	Advanced cancer	Survival prognosis and treatment of patients with colorectal cancer
Vadhwana et al. [26]	–	Microscopic features in histology
Prabhakaran et al. [27]	–	Survival prognosis and treatment of patients with rectal cancer

covering databases, algorithms, and metrics applied in CI for colorectal cancer.

4.4.1. Data set

Table 5 lists datasets containing colonoscopy and histology images: CVC-ColonDB, CVC-ClinicDB, ETIS-LARIB, ASU-Mayo, CVideoClinicDB, Kvasir, Nerthus, CVC-Endo- SceneStill, and Mesejo. The Warwick QU database includes tissue images relevant for microscopic analysis of lower gastrointestinal polyps. Other electronic health record sources like THIN and Utrecht Primary Care dataset were noted but excluded due to non-colorectal cancer focus.

Table 6 details various datasets, each with data type (histology or colonoscopy images), size (number of images), and imaging equipment (white-light or narrow-band colonoscopes, Zeiss MIRAX MIDI). Examples include CVC-ColonDB (300 white-light colonoscopy images),

Table 5
Characteristics of the extracted datasets.

Reference	Dataset
Thakur et al. [7]	Warwick QU dataset
Davri et al. [20]	Warwick QU dataset
Alam et al. [21]	Warwick QU dataset
Sanchez-Peralta et al. [9]	CVC-ColonDB; CVC-ClinicDB; ETIS-LARIB; ASU-Mayo; CVCVideoClinicDB; Kvasir; Nerthus; CVC-EndoSceneStill; Mesejo
Nazarian et al., 2021	CVC-ColonDB; ASU-Mayo

Table 6
Datasets with colonoscopy and histology images extracted from selected literature reviews.

Dataset	Data types	Data set size	Obtaining instrument
Warwick QU dataset	Images of histological slides	165 images	Zeiss MIRAX MIDI
CVC-ColonDB	Colonoscopy pictures and videos	300 images	White-Light Colonoscopy
CVC-ClinicDB	Colonoscopy pictures and videos	612 images	White-Light Colonoscopy
ETIS-LARIB	Colonoscopy pictures and videos	196 images	White-Light Colonoscopy
ASU-Mayo	Colonoscopy pictures and videos	38 small SD and HD video	Narrow Band Imaging and White-Light Colonoscopy
Kvasir	Colonoscopy pictures and videos	1 000 polyp images	–
EndoScene- Nerthus	Colonoscopy pictures and videos	2.597 images	–
CVC-EndoSceneStill	Colonoscopy pictures and videos	912 WL images extracted from 44 videos.	CVC-ClinicDB + CVC-ColonDB
Mesejo	Colonoscopy pictures and videos	76 short videos	Narrow Band Imaging and White-Light Colonoscopy

Kvasir (1000 colorectal polyp images), ASU-Mayo, Mesejo (narrow-band colonoscopy), and Warwick QU (both colonoscopy and histology).

4.4.2. Algorithms

Most of the selected studies (18 out of 22) focus on machine learning algorithms, and four of them address performance metrics in prediction models [6,12,25,26]. Fig. 3(a) highlights prominent algorithms like Convolutional Neural Networks (CNN) used in polyp detection and classification studies, along with Decision Trees (DT), Gradient Boosting (GB), Random Forest (RF), Classification and Regression Tree, and Naive Bayes.

AI applications in histological analysis aim to detect suspicious colorectal cancer cases for enhanced diagnosis and treatment. Algorithms such as Two path CNN, U-Net, FCN, CNN, K-means/naıve Bayesian, AlexNet, ResNet, VGG, RNN, ResNet18, LSTM, Separator-Net, Smoothness, Topology, and Holistically offer various advantages tailored to specific tasks. For colorectal cancer detection and screening, CNN, SVM, Fourier Filters, GrayRLM, ANN, Nearest Neighbor, MLP, Naive Bayes, and Random Forest algorithms are utilized, facilitating early disease identification and more effective treatment strategies.

4.4.3. Evaluation metrics

The Fig. 3(b) displays the frequency of evaluation metrics across 22 selected articles. Most studies (20) aimed to assess CI algorithms for diagnosing and treating colorectal cancer, with one study [22] focusing on CI algorithms in statistical models for detecting the disease's incidence. Sensitivity and specificity were the most frequently used metrics, appearing in 16 studies. Overall, the models in the reviewed studies primarily addressed colorectal cancer diagnosis and screening, except for [15], which focused on prognostic data tracking.

4.4.4. Matrix of article correlation

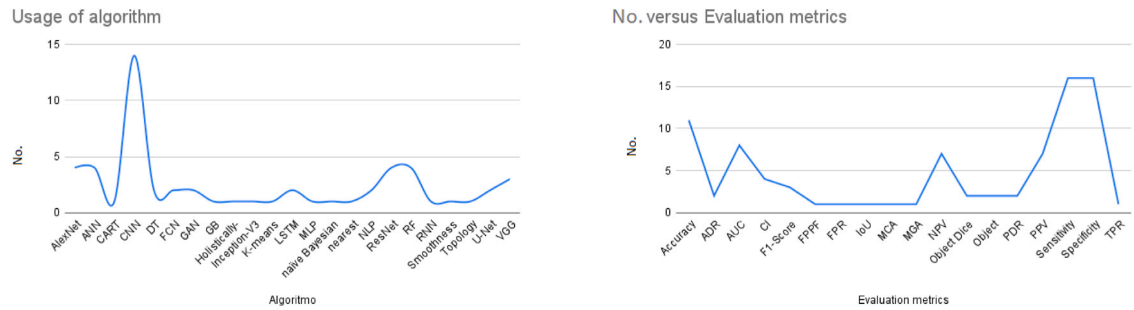
To evaluate the overlap among the studies identified in the literature reviews, Fig. 4 presents a graph of the main primary studies mentioned. We identified 52 articles cited in two or more reviews. Seven studies showed no correlation: Liang et al. [15]; Moen et al. [19]; Alsadhan et al. [22]; Seow-En et al. [24]; Wang et al. [25]; Vadhwana et al. [26]. Overlap was classified into four categories, with most studies showing low to moderate overlap. Of the 54 studies, 35 had overlap below 15%, 11 had moderate overlap: Byrne et al. [30]; Renner et al. [31]; Shin and Balasingham [32]; Yamada et al. [33]; Bychkov et al. [34]; Wang et al. [35]; Repici et al. [36]; Sirinukunwattana et al. [37]; Xu et al. [38]; Kainz et al. [39]; Graham et al. [40]. Six showed high overlap: Ozawa et al. [41]; Chen et al. [42]; Kather et al. [43]; Wang et al. [44]; Liu et al. [45]; Su et al. [46]. Two showed very high overlap: Urban et al. [47]; Wang et al. [48].

4.4.5. Risk of bias in studies — reporting biases of the systematic reviews

The AMSTAR 2 evaluation rigorously assesses systematic reviews, offering critical insights for clinical decision-making and future research directions. Five reviews were classified as high quality, exemplifying strong methodological rigor [6,7,13,16,21,23]. Another eight reviews were rated moderate, indicating areas for enhancement [8–12, 14,19,20,24–27]. Four reviews were deemed low quality, highlighting methodological deficiencies [15,17,18,22]. These findings emphasize the importance of considering review quality when interpreting results, ensuring robust foundations for clinical decisions and future research in CI for colorectal cancer.

5. Discussion

As mentioned in the introduction, this study addresses a gap not yet explored in the literature by investigating systematic reviews on



(a) ANN (artificial neural network); CART (classification and regression tree); CNN (convolutional neural network); DT(decision trees); FCN (fully convolutional network); GAN (generative adversarial network); GB (gradient boosting); LSTM (long short term memory); MLP (Multilayer perceptron); NLP (natural language processing); RF(random forests); RNN (recurrent neural networks)

(b) ADR (Adenoma Detection Rate); AUC (Area Under the Curve); CI (Confidence Interval); FPPF (False Positive Per Frame); FPR (False Positive Rate); IoU (Intersection over Union); MCA (Mean Class Accuracy); MGA (Mean Global Accuracy); NPV (Negative Predictive Value); PDR (Polyp Detection Rate); PPV (Positive Predictive Value); TPR (True Positive Rate)

Fig. 3. Algorithms and metrics reported from the analyzed reviews.

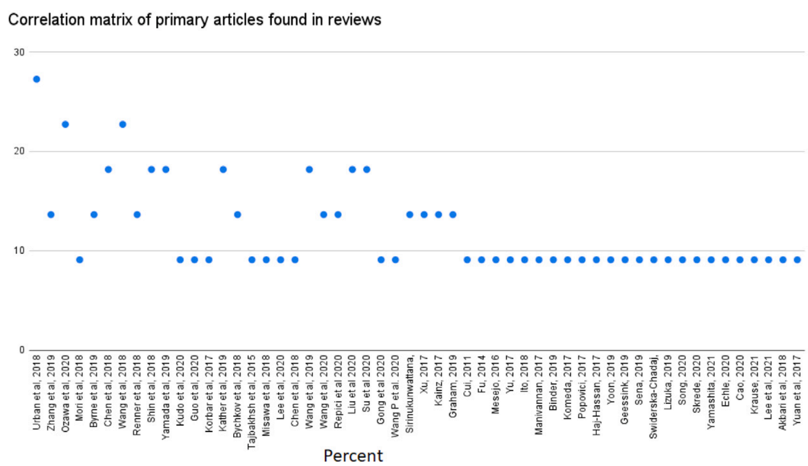


Fig. 4. Matrix of article correlation.

the use of machine learning in colorectal cancer— a topic still unexamined by other tertiary studies. This originality reinforces its relevance by organizing and synthesizing evidence in an emerging field with high impact on global health. This section discusses the contributions of computational intelligence (CI) to the diagnosis and treatment of colorectal cancer, based on the reviews analyzed and the research questions associated with the scopes described in Section 4.3 and Table 4.

5.1. *Microscopic features in histology*

What are the main contributions of computational intelligence to histological analysis? The studies analyzed indicate that CI plays an essential role in reducing intra- and inter-observer variability and in interpreting cellular heterogeneity. Deep learning models, particularly convolutional neural networks (CNNs), have been widely used in the classification of histological images with high accuracy, assisting therapeutic decision-making [7,13,21]. In addition, advanced CI techniques have enabled the identification of molecular patterns, such as microsatellite instability (MSI-H) and features of the tumor microenvironment (TME), both of which are relevant prognostic factors in precision oncology [20,21]. Computer-aided detection (CAD) systems have also been progressively incorporated into clinical practice, contributing to greater diagnostic reproducibility and reduced subjectivity among specialists [26].

What advancements are needed to overcome the limitations of histological image datasets used in algorithm training? The studies emphasize the need for larger, more representative datasets that include ethnic and regional diversity [13]. The lack of standardized annotations compromises reproducibility, highlighting the urgent need for stricter guidelines [7]. Despite progress in CNN-based models, the scarcity of multiethnic datasets and their still inferior performance compared to human specialists point to the importance of protocols such as STARD-AI to ensure validation and safe implementation in clinical practice [20,21,26].

5.2. *Macroscopic features at colonoscopy*

Macroscopic features observed during colonoscopy play an essential role in the identification and classification of colonic lesions. This section discusses how computational intelligence (CI) can enhance real-time classification and address lesion variability.

What are the current limitations in using CI for real-time classification of macroscopic lesions? CI models tend to perform worse in detecting flat lesions, which are more visually subtle and harder to recognize. Furthermore, high false-positive rates and performance variations based on the operator's experience remain significant limitations. Yamada et al. [12] demonstrated that these models achieve better results with elevated lesions than with superficial ones.

How can the integration of algorithms help manage variability in lesion presentation? Some studies suggest combining detection, localization, and segmentation approaches using deep learning, along with techniques like data augmentation and patch-based analysis [48, 49]. Nevertheless, there is no consensus on which deep learning architecture is most robust when dealing with the morphological variability of colonic lesions [9].

How do the reviews discuss the influence of dataset size and quality on the performance of CI techniques for macroscopic classification? Many reviews point out that existing colonoscopy datasets are often small, heterogeneous, and proprietary, which limits reproducibility and prevents fair comparisons across methods. The lack of standardized, multicenter datasets hinders external validation and broad clinical applicability of CI models [10].

5.3. Stage of the disease

Staging colorectal cancer is fundamental for guiding therapeutic planning and assessing disease progression. This section explores the contribution of CI to staging accuracy and the associated implementation challenges.

How can artificial intelligence improve the accuracy of lymph node metastasis staging? AI models, especially those based on deep learning, can identify subtle features in CT and MRI scans that might be missed by radiologists. These models have shown higher performance levels, with AUROCs reaching up to 0.917, surpassing human evaluation benchmarks (AUROC 0.688) [6,17].

5.4. Survival prognosis and treatment of patients with colorectal cancer

This section addresses how CI can improve prognostic accuracy and treatment planning in colorectal cancer.

How can CI more accurately predict survival and treatment response? Computational models such as convolutional neural networks (CNNs), support vector machines (SVMs), and random forests (RFs) can capture complex, nonlinear interactions among clinical, pathological, and radiomic data. Hybrid models that integrate clinical and imaging features have achieved high predictive performance, with AUCs of up to 0.870 for liver metastasis prediction [24], contributing to more personalized prognostic assessments.

What are the ethical and technical challenges in integrating CI into clinical prognosis? Key technical limitations include the lack of external validation, insufficient data standardization, and limited generalizability across institutions. Ethically, the opacity of many AI models – the so-called "black-box" problem – hinders clinical trust. Additionally, algorithmic bias arising from imbalanced datasets may disproportionately impact underrepresented populations. To ensure safety and fairness, transparency, auditability, and interpretability must be prioritized [15].

5.5. Cancer screening and surveillance

This section addresses the following research question: **What barriers must be overcome for the large-scale adoption of computational intelligence (CI) in screening and surveillance?** According to [16], many AI models are trained using idealized images and do not reflect real-world challenges, such as bleeding, mucus, and low image quality. Moreover, the lack of validation in real clinical settings, algorithmic inconsistencies, and the scarcity of prospective clinical trials hinder the standardization and practical applicability of these systems. Overcoming these obstacles is essential to ensure the reliable use of CI in population-level screening and surveillance strategies.

5.6. Implications for clinical translation and research

The findings of this review highlight important implications for various stakeholders involved in the development and implementation of computational intelligence (CI) in colorectal cancer (CRC) care.

Healthcare professionals should exercise caution when adopting unvalidated models, as most of the reviews analyzed reported a lack of external validation and reliance on non-representative datasets, limiting the clinical applicability of current CI tools [4,16]. Integrating these tools into routine practice requires robust validation in real-world settings and diverse patient populations.

Researchers should prioritize methodological standardization, including data collection, annotation, and evaluation protocols, as well as the use of consistent performance metrics to enable cross-study comparisons. Addressing dataset imbalance and improving reproducibility are also critical challenges to be addressed.

Funding agencies and policymakers should support interdisciplinary research and the development of public, balanced, and diverse datasets. This infrastructure is essential to ensure fairness and accelerate the responsible adoption of CI in healthcare systems.

Aligning technological development with clinical needs and ethical principles will enable the transition from experimental models to safe, effective, and scalable solutions in CRC management.

5.7. Future directions

Future studies on computational intelligence (CI) in colorectal cancer (CRC) should prioritize external validation across diverse, multicenter datasets to improve model generalizability. The development of interpretable models using tools like SHAP and LIME is essential to increase transparency and clinical trust.

Standardized protocols for image acquisition, annotation, and model evaluation are needed to reduce methodological variability. Public datasets should be expanded to address class imbalance and include greater population diversity.

There is also a need to integrate multimodal data – such as imaging, histology, and electronic health records – to support more accurate prognosis and treatment planning. Research should move beyond performance metrics and assess clinical utility through prospective studies and real-world validation.

Finally, ethical aspects must be considered, especially regarding algorithmic bias and its impact on underrepresented groups. Interdisciplinary collaboration will be key to translating CI advances into effective and safe tools for CRC care.

5.8. Contributions

This tertiary review provides an unprecedented synthesis of the current evidence on the use of CI in CRC, filling a relevant gap in the literature. It includes a detailed mapping of the algorithms, metrics, and datasets used in the primary studies referenced by the analyzed reviews, as well as an overlap analysis using a correlation matrix to highlight influential studies and unaddressed gaps. Methodological quality was assessed using the AMSTAR 2 tool, revealing both strengths and weaknesses in the current body of literature. The classification of studies into clinically relevant scopes – such as histology, colonoscopy, staging, prognosis, and screening – helps bridge technological advances with medical decision-making. Based on these findings, practical recommendations were proposed to guide future research and support the safe integration of CI into clinical practice.

5.9. Limitations

Several limitations must be considered. This study was conducted as an expanded narrative review, although guided by established protocols such as PRISMA, and the search was restricted to the PubMed and EMBASE databases, which may have excluded relevant studies available in other sources. Methodological heterogeneity among the included reviews – such as differences in selection criteria and bias assessment tools – hinders direct comparison across studies. Furthermore, no meta-analysis or reassessment of primary data was performed, which limits the statistical robustness of the synthesis.

6. Conclusions

This review offers a comprehensive overview of recent advancements in computational intelligence applied to colorectal cancer, covering its use in detection, diagnosis, staging, and prognosis. A total of 22 systematic reviews were analyzed, highlighting the transformative potential of CI in clinical care. However, several challenges remain, including limited and non-diverse datasets, lack of standardized protocols, and the use of black-box models. These limitations underscore the need for interdisciplinary collaboration to develop more robust, interpretable, and clinically applicable solutions. The findings from this review provide a valuable foundation for future research and the evolution of AI-based approaches in colorectal cancer care.

CRediT authorship contribution statement

Aurea Valeria Pereira Silva: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Plinio Sa Leitao-Junior:** Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Ethical statement

We, Aurea Valeria Pereira Silva and Plinio Sa Leitao-Junior, declare that the article entitled "Advancements in Artificial Intelligence for Colorectal Cancer: A Comprehensive Overview of Systematic Reviews" was conducted in accordance with ethical principles, maintaining scientific integrity and responsible research conduct.

The study consists of a literature review and did not involve any direct experimentation with humans or animals. All sources used were properly cited, and there was no form of plagiarism or scientific misconduct.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Authors would like to thank CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil), financial code #001.

References

[1] Collins T, Maktabi M, Barberio M, Bencteux V, Jansen-Winkeln B, Chalopin C, et al. Automatic recognition of colon and esophagogastric cancer with machine learning and hyperspectral imaging. Diagnostics 2021;11(10):1810.
[2] INCA. Estimativa 2020 : Incidência de câncer no Brasil. Rio de Janeiro: INCA; 2019.
[3] Stark UA, Kohlmann K, Buchner S, Jordan J, Siebenhofer A, Raspe H. What is the effectiveness of various invitation methods to a colonoscopy in the early detection and prevention of colorectal cancer? Protocol of a systematic review. Syst Rev 2020;9(1):1–7.
[4] Cao J-S, Zhao X-Q, Sun Y-S, Li X-L. Artificial intelligence in gastroenterology and hepatology: Status and challenges. World J Gastroenterol 2021;27(16):1664.
[5] Selçuk AA. A guide for systematic reviews: PRISMA. Turk Arch Otorhinolaryngol 2019;57(1):57.
[6] Bedrikovetski S, Dudi-Venkata NN, Kroon HM, Seow W, Vather R, Carneiro G, et al. Artificial intelligence for pre-operative lymph node staging in colorectal cancer: a systematic review and meta-analysis. BMC Cancer 2021;21:1–10.
[7] Thakur N, Yoon H, Chong Y. Current trends of artificial intelligence for colorectal cancer pathology image analysis: a systematic review. Cancers 2020;12(7):1884.
[8] Xu Y, Ding W, Wang Y, Tan Y, Xi C, Ye N, et al. Comparison of diagnostic performance between convolutional neural networks and human endoscopists for diagnosis of colorectal polyp: A systematic review and meta-analysis. PLoS One 2021;16(2):e0246892.
[9] Sanchez-Peralta LF, Bote-Curiel L, Picon A, Sanchez-Margallo FM, Pagador JB. Deep learning to find colorectal polyps in colonoscopy: A systematic literature review. Artif Intell Med 2020;108:101923.
[10] Nazarian S, Glover B, Ashrafian H, Darzi A, Teare J. Diagnostic accuracy of artificial intelligence and computer-aided diagnosis for the detection and characterization of colorectal polyps: systematic review and meta-analysis. J Med Internet Res 2021;23(7):e27370.
[11] 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(1):77–85.
[12] Yamada M, Saito Y, Yamada S, Kondo H, Hamamoto R. Detection of flat colorectal neoplasia by artificial intelligence: A systematic review. Best Pr Res Clin Gastroenterol 2021;52:101745.
[13] Kuntz S, Krieghoff-Henning E, Kather JN, Jutzi T, Höhn J, Kiehl L, et al. Gastrointestinal cancer classification and prognostication from histology using deep learning: Systematic review. Eur J Cancer 2021;155:200–15.
[14] Aziz M, Fatima R, Dong C, Lee-Smith W, Nawras A. The impact of deep convolutional neural network-based artificial intelligence on colonoscopy outcomes: a systematic review with meta-analysis. J Gastroenterol Hepatol 2020;35(10):1676–83.
[15] Liang H, Yang L, Tao L, Shi L, Yang W, Bai J, et al. Data mining-based model and risk prediction of colorectal cancer by using secondary health data: a systematic review. Chin J Cancer Res 2020;32(2):242.
[16] Mohan BP, Khan SR, Kassab LL, Ponnada S, Dulai PS, Kochhar GS. Accuracy of convolutional neural network-based artificial intelligence in diagnosis of gastrointestinal lesions based on endoscopic images: A systematic review and meta-analysis. Endosc Int Open 2020;8(11):E1584–94.
[17] Moreira JM, Santiago I, Santinha J, Figueiredo N, Marias K, Figueiredo M, Vanneschi L, Papanikolaou N. Challenges and promises of radiomics for rectal cancer. Curr Color Cancer Rep 2019;15:175–80.
[18] Langarizadeh M, et al. Machine learning techniques for diagnosis of lower gastrointestinal cancer: a systematic review. Iran. Red Crescent Med. J. (IRCMJ) 2021.
[19] Moen S, Vuik FE, Kuipers EJ, Spaander MC. Artificial intelligence in colon capsule endoscopy—A systematic review. Diagnostics 2022;12(8):1994.
[20] Davri A, Birbas E, Kanavos T, Ntritsos G, Giannakeas N, Tzallas AT, et al. Deep learning on histopathological images for colorectal cancer diagnosis: A systematic review. Diagnostics 2022;12(4):837.
[21] Alam MR, Abdul-Ghafar J, Yim K, Thakur N, Lee SH, Jang H-J, et al. Recent applications of artificial intelligence from histopathologic image-based prediction of microsatellite instability in solid cancers: a systematic review. Cancers 2022;14(11):2590.
[22] Alsadhan N, Almaiman A, Pujades-Rodriguez M, Brennan C, Shuweihdi F, Alhurishi SA, et al. Statistical methods for measuring trends in colorectal cancer incidence in registries: A systematic review. Front Oncol 2022;12.
[23] Keshtkar K, Safarpour AR, Heshmat R, Sotoudehmanesh R, Keshtkar A. A systematic review and meta-analysis of convolutional neural network in the diagnosis of colorectal polyps and cancer. Turk J Gastroenterol 2023;34(10):985.
[24] Seow-En I, Koh YX, Zhao Y, Ang BH, Tan I, Chok AY, et al. Predictive modeling algorithms for liver metastasis in colorectal cancer: A systematic review of the current literature. Ann Hepato-biliary-Pancreat Surg 2023.
[25] Wang Q, Xu J, Wang A, Chen Y, Wang T, Chen D, et al. Systematic review of machine learning-based radiomics approach for predicting microsatellite instability status in colorectal cancer. La Radiol Medica 2023;128(2):136–48.

[26] Vadhwana B, Tarazi M, Patel V. The role of artificial intelligence in prospective real-time histological prediction of colorectal lesions during colonoscopy: A systematic review and meta-analysis. Diagnostics 2023;13(20):3267.

[27] Prabhakaran S, Choong KWK, Prabhakaran S, Choy KT, Kong JC. Accuracy of deep neural learning models in the imaging prediction of pathological complete response after neoadjuvant chemoradiotherapy for locally advanced rectal cancer: a systematic review. Langenbeck' s Arch Surg 2023;408(1):321.

[28] Short MW, Layton SC, Teer JK. Colorectal cancer screening and surveillance. Am Fam Physician 2015;91(2):93–100.

[29] Wang Y-P, Guo P-T, Zhu Z, Zhang H, Xu Y, Ma S-P, et al. Macroscopic serosal classification of colorectal cancer and its clinical significance. Int J Clin Exp Med 2015;8(11):20123.

[30] Byrne MF, Chapados N, Soudan F, Oertel C, Pérez ML, 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(1):94–100.

[31] Renner J, Phlipsen H, Haller B, Navarro-Avila F, Saint-Hill-Febles Y, Mateus D, et al. Optical classification of neoplastic colorectal polyps–a computer-assisted approach (the COACH study). Scand J Gastroenterol 2018;53(9):1100–6.

[32] Shin Y, Balasingham I. Automatic polyp frame screening using patch based combined feature and dictionary learning. Comput Med Imaging Graph 2018;69:33–42.

[33] Yamada M, Saito Y, Imaoka H, Saiko M, Yamada S, Kondo H, et al. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. Sci Rep 2019;9(1):14465.

[34] Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. Sci Rep 2018;8(1):3395.

[35] Wang P, Liu X, Berzin TM, Brown JRG, Liu P, Zhou C, et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. Lancet Gastroenterol Hepatol 2020;5(4):343–51.

[36] 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(2):512–20.

[37] Sirinukunwattana K, Pluim JP, Chen H, Qi X, Heng P-A, Guo YB, et al. Gland segmentation in colon histology images: The glas challenge contest. Med Image Anal 2017;35:489–502.

[38] Xu Y, Jia Z, Wang L-B, Ai Y, Zhang F, Lai M, et al. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. BMC Bioinformatics 2017;18(1):1–17.

[39] Kainz P, Pfeiffer M, Urschler M. Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. PeerJ 2017;5:e3874.

[40] Graham S, Chen H, Gamper J, Dou Q, Heng P-A, Snead D, et al. MILD-net: Minimal information loss dilated network for gland instance segmentation in colon histology images. Med Image Anal 2019;52:199–211.

[41] Ozawa T, Ishihara S, Fujishiro M, Kumagai Y, Shichijo S, Tada T. Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks. Ther Adv Gastroenterol 2020;13:1756284820910659.

[42] Chen P-J, Lin M-C, Lai M-J, Lin J-C, Lu HH-S, Tseng VS. Accurate classification of diminutive colorectal polyps using computer-aided analysis. Gastroenterology 2018;154(3):568–75.

[43] Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nature Med 2019;25(7):1054–6.

[44] Wang P, Berzin TM, Brown JRG, Bharadwaj S, Becq A, Xiao X, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut 2019;68(10):1813–9.

[45] Liu W-N, Zhang Y-Y, Bian X-Q, Wang L-J, Yang Q, Zhang X-D, et al. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. Saudi J Gastroenterol: Off J the Saudi Gastroenterol Assoc 2020;26(1):13.

[46] Su J-R, Li Z, Shao X-J, Ji C-R, Ji R, Zhou R-C, et al. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). Gastrointest Endosc 2020;91(2):415–24.

[47] Urban G, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology 2018;155(4):1069–78.

[48] Wang P, Xiao X, Glissen Brown JR, Berzin TM, Tu M, Xiong F, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng 2018;2(10):741–8.

[49] Zhang M, Li H, Li Y, Zhang M, Zou W. Differential diagnosis for esophageal protruded lesions using a deep convolution neural network in endoscopic images. Gastrointest Endosc 2021;93(6):1261–72.e2.