IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 28 May 2023, accepted 26 June 2023, date of publication 30 June 2023, date of current version 7 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3290997

TOPICAL REVIEW

Recent Advances in Applying Machine Learning and Deep Learning to Detect Upper Gastrointestinal Tract Lesions

MALINDA VANIA[1,2], BAYU ADHI TAMA[3], (Member, IEEE), HASAN MAULAHELA[4], AND SUNGHOON LIM[1,2,5]

[1]Department of Industrial Engineering, Ulsan National Institute of Science and Technology, Ulju-gun, Ulsan 44919, Republic of Korea
[2]Industry Intelligentization Institute, Ulsan National Institute of Science and Technology, Ulju-gun, Ulsan 44919, Republic of Korea
[3]NSF HDR Institute for Harnessing Data and Model Revolution in the Polar Regions (iHARP), University of Maryland, Baltimore County (UMBC), Baltimore, MD 21250, USA
[4]Division of Gastroenterology, Department of Internal Medicine, Faculty of Medicine, Universitas Indonesia-Cipto Mangunkusumo National General Hospital, Jakarta 10430, Indonesia
[5]Graduate School of Artificial Intelligence, Ulsan National Institute of Science and Technology, Ulju-gun, Ulsan 44919, Republic of Korea

Corresponding author: Sunghoon Lim (sunghoonlim@unist.ac.kr)

This work was supported in part by the Advanced Technology Center Plus (ATC+) Program, 50% Accident Prevention Focus to Reduce Accident Rate Development of Risk Detection System for Road Facilities Based on Artificial Intelligence, funded by the Ministry of Trade, Industry and Energy (MOTIE), under Grant 20017932; and in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government [Ministry of Science and ICT (MSIT)] under Grant 2021R1F1A1046416.

ABSTRACT The clinical application of a real-time artificial intelligence (AI) image processing system to diagnose upper gastrointestinal (GI) malignancies remains an experimental research and engineering problem. Understanding these commonly used technical techniques is required to appreciate the scientific quality and novelty of AI studies. Clinicians frequently lack this technical background, and AI experts may be unaware of such clinical relevance and implications in daily practice. As a result, there is a growing need for a multidisciplinary, international assessment of how to conduct high-quality AI research in upper GI malignancy detection. This research will help endoscopists build approaches or models to increase diagnosis accuracy for upper GI malignancies despite variances in experience, education, personnel, and resources, as it offers real-time and retrospective chances to improve upper GI malignancy diagnosis and screening. This comprehensive review sheds light on potential enhancements to computer-aided diagnostic (CAD) systems for GI endoscopy. The survey includes 65 studies on automatic upper GI malignancy diagnosis and evaluation, which are compared by endoscopic modalities, image counts, models, validation methods, and results. The main goal of this research is to assess and compare each AI method's current stage and potential improvement to boost performance, maturity, and the possibility to open new research areas for the application of a real-time AI image recognition system that diagnoses upper GI malignancies. The findings of this study suggest that Support Vector Machines (SVM) are frequently utilized in gastrointestinal (GI) image processing within the context of machine learning (ML). Moreover, the analysis reveals that CNN-based supervised learning object detection models are widely employed in GI image analysis within the deep learning (DL) context. The results of this study also suggest that RGB is the most commonly used image modality for GI analysis, with color playing a vital role in detecting bleeding locations. Researchers rely on public datasets from 2018-2019 to develop AI systems, but combining them is challenging due to their unique classes. To overcome the problem of insufficient data to train a new DL model, a standardized database is needed to hold different datasets for the development of AI-based GI endoscopy systems.

The associate editor coordinating the review of this manuscript and approving it for publication was Gustavo Olague.

INDEX TERMS Endoscopy, upper gastrointestinal tract, artificial intelligence, machine learning, deep learning.

LIST OF ABBREVIATIONS

Acc	Accuracy.
AI	Artificial Intelligence.
AUC	Area Under Curve.
BE	Barrett's Esophagus.
CAD	Computer-Aided Diagnostic.
CAG	Chronic Atrophic Gastritis.
CLBP	Complete Local Binary Pattern.
CNN	Convolutional Neural Network.
DL	Deep Learning.
DSC	DICE Score.
EAC	Esophageal Adeno-Carcinoma.
ESCC	Esophageal Squamous Cell Carcinoma.
F1	F1 Score.
FL	Fluorescence Endoscopy.
FPR	False Positive Ratio.
GA	Genetic Algorithm.
GAC	Gastric Adeno-Carcinoma.
GI	Gastrointestinal.
GIM	Gastric Intestinal Metaplasia.
HAF	Hybrid Adaptive Filtering.
HC	Handcrafted.
Hp	Helicobacter Pylori.
IM	Intestinal Metaplasia.
IoU	Intersection Over Union.
KAP	Cohen's Kappa Score.
KNN	K-Nearest Neighbor.
LLC	Locality-Constrained Linear Coding.
ML	Machine Learning.
MLP	Multi-Layer Perceptron.
NBI	Narrow Band Imaging.
NPV	Negative Predictive Value.
PPV	Positive Predictive Value.
Prec	Precision.
R	Recall.
Sen	Sensitivity.
Spec	Specificity.
SSD	Single Shot Multibox Detection.
SVM	Support Vector Machine.
UC	Ulcerative Colitis.
WCE	Wireless Capsule Endoscopy.
WLE	White Light Endoscopy.
WLI	White Light Imaging.

I. INTRODUCTION

The most prevalent cancers found throughout the world are those that affect the upper gastrointestinal (GI) system. These forms of cancers include esophageal and gastric malignancies. In contrast to colon polyps, the morphology of upper GI abnormalities can vary greatly, which presents a problem for upper GI endoscopy. In addition, the pathophysiology is distinct in Asia, Europe, and North America. For instance, neoplasia associated with Barrett's esophagus is more common in Europe and North America than in Asia, where squamous neoplasia is more common. Despite these differences, the development of artificial intelligence (AI) for real-time GI endoscopy has made significant progress, and this review article represents a significant step forward in that evolution.

Although AI platforms powered by deep learning (DL) algorithms are making tremendous progress in the medical imaging field, there has been little success in applying them to diagnose upper GI malignancies. Therefore, while AI application in medicine is causing significant excitement within the medical community, there is also ambiguity and worry surrounding this topic. Researchers seek a global uniform methodology for creating and validating such AI systems, ensuring various peoples and diseases are represented during the design, development, validation, and testing phases of a GI AI diagnostic system. It is, therefore, necessary to assist individuals during the design and validation process of a GI AI diagnostic system. This system will use machine learning (ML) and DL to analyze images obtained from clinical endoscopies to identify upper GI malignancies. The contributions of these systems should hold the confidence of both medical practitioners and patients.

The information presented in this paper is expected to assist further non-expert endoscopists who practice in primary, essential, or low-volume hospitals in developing a method or model to achieve diagnostic accuracy for upper GI malignancies comparable to that of expert endoscopists. This paper will provide both real-time and retrospective opportunities for all providers to improve the efficiency of diagnosing and screening for upper GI malignancies. This work includes a comprehensive survey that discusses advances in computer-aided diagnosis (CAD) systems in several endoscopy modalities that are used for GI examination. These endoscopy modalities include the following: (I) white light endoscopy (WLE), (II) high-definition white light endoscopy (HD-WLE), and (III) narrow band imaging (NBI). The survey comprises 65 papers on the automatic detection and evaluation of upper GI malignancies. These papers are compared by their endoscopy modalities, number of images, and models applied to the problems, validation methods, and results. The primary objective is to provide a comparative analysis of present methods to open avenues for future research in this field.

The database was searched using the terms "artificial intelligence", "AI", "machine learning", "deep learning", "capsule endoscopy", "upper gastrointestinal", "capsule endoscope", "machine intelligence", "computational intelligence", "image recognition", and "convolutional neural network" as well as similar terms for capsule endoscopy, such as "wireless capsule endoscopy", "video capsule endoscopy", "gastrointestinal diagnosis", "esophageal and gastroduodenal endoscopy'', ''neoplasia'', ''Barrett's esophagus'', ''esophagitis'', ''gastro-esophageal'', ''ulcer detection'', ''bleeding detection'', ''gastric cancer'', ''Helicobacter pylori infection'', ''upper gastrointestinal cancer'', ''esophageal cancer'', and ''gastric cancer'' to find publications on AI-based endoscopy diagnostic systems for upper GI malignancies from January 1, 2015 to March 31, 2022. The search for information in the scientific literature yields few results. Due to the retrospective nature of the studies, their concentration on a single disease, their small sample size, and their lack of validation, definitive conclusions regarding the clinical applicability and reliability of an AI-assisted endoscopic diagnosis remain uncertain. This is the case even though there have been some encouraging preliminary reports. The clinical application of a real-time AI image recognition system to diagnose upper GI malignancies has thought to be experimental up until this point.

The focus of this survey is to give a comprehensive overview of the approaches that have been recently developed and utilized for identifying upper GI malignancies, such as tumors, polyps, and ulcers, using wireless capsule endoscopy (WCE) as the only source. In particular, generalized anomalies detected in WCE images, primarily bleeding/lesion detection, are presented to circumvent the knowledge constraints. Bleeding is one of the most critical problems that should be detected in WCE, because it is one of the most common complications associated with WCE [1]. Because bleeding in the GI tract is typically the result of one or more illnesses, using WCE images to diagnose various types of GI bleeding has recently attracted substantial attention. This issue is the subject of a significant number of research articles [2], [3], [4], [5], [6], [7], [8], [9], [10]. However, there are new hurdles associated with WCE technology for clinicians to overcome in order to detect bleeding in patients. Given that the WCE creates 55,000 images for each examination, it is highly laborious for doctors to manually scan these images frame by frame to identify and discover those that show bleeding. In addition, the images that capture GI tract anomalies make up fewer than five percent of the total number of images the WCE collects. Additionally, some bleeding regions might not be visible to the human eye at all because of inconsistent lighting. As a result, it is ultimately essential to develop an automatic computer-aided system that assists doctors in detecting and analyzing bleeding images. In the following sections, the technique utilized in the evaluation of the cited articles as well as some medical background information pertaining to the disease, will be presented.

The remaining parts of the paper are structured as described below. In Section II, the problem definitions and motivation for abnormality identification in WCE images are discussed. In Section III, the overall process used for this systematic mapping study is illustrated. Following this process, Section IV summarizes a selection of recent research articles that are pertinent to WCE imaging. This section will also attempt to explain and address numerous concerns that have arisen regarding the use of ML and DL approaches in GI image analysis within the literature that has been cited in this review. A tabular form is used to display information regarding each method's performance for the detection of relevant anomalies. In Section V, the findings of the empirical benchmark used to evaluate current state-of-the-art DL models for GI image processing are presented. In addition, this section offers a summary and conclusion of the presented survey and identifies potential areas for improvement.

II. PROBLEM DEFINITIONS AND MOTIVATIONS

Numerous abnormalities in the mucosal lining of the GI tract, ranging from minor irritations to life-threatening conditions, can occur. According to the World Health Organization (WHO)'s specialized cancer agency, the International Agency for Research on Cancer, GI cancers account for approximately 3.6 million new cases worldwide each year [11], and 1.6 million cases are related to stomach and esophagus cancers. Approximately 2.7 million people die yearly from these cancers, with around 1.3 million deaths related to stomach and esophagus cancers (Figure 1).

Due to technical advances in healthcare practices and greater research access to vast medical databases, the diagnosis and treatment of many diseases have improved. The incorporation of new technologies into clinical practice may be a crucial component. Early identification of GI cancers could greatly enhance survival rates to as high as 90% [12]. Endoscopy is the most effective method for detecting and diagnosing GI cancers. However, the detection accuracy depends on the skill of the endoscopists and is hindered by a variety of GI variables. This investigation is deemed crucial for reducing GI cancer incidences and deaths.

After each endoscopy, gastroenterologists should produce endoscopic procedure reports, which are an important part of their work. Minimal Standard for Reporting (MSR) and Minimal Standard Terminology (MST) are recommended by the World Endoscopy Organization (WEO) [13]. Although endoscopy is now the gold standard for inspecting the GI tract, the diversity in operator performance significantly limits its effectiveness. In order to avoid GI disease-related morbidity and death, enhanced endoscopic performance, high-quality clinical examinations, and systematic screening are essential. As AI-enabled support systems have emerged, they have shown promise in providing healthcare personnel with the tools they need to offer high-quality care on an industrial-scale basis.

AI has recently garnered significant interest in various medical sectors. The advancement of AI techniques, such as ML and DL, has expanded medical imaging analysis capabilities. Endoscopists have long focused on preventing GI cancers by screening endoscopy, and they have recently turned their attention to AI applications in capsule endoscopy. As a result, newer endoscopic procedures, like capsule endoscopy, provide advantages that may overcome some limitations of standard upper GI endoscopy. Capsule endoscopy is a painless, non-invasive treatment that is widely used. However, wireless capsule endoscopy analysis is a

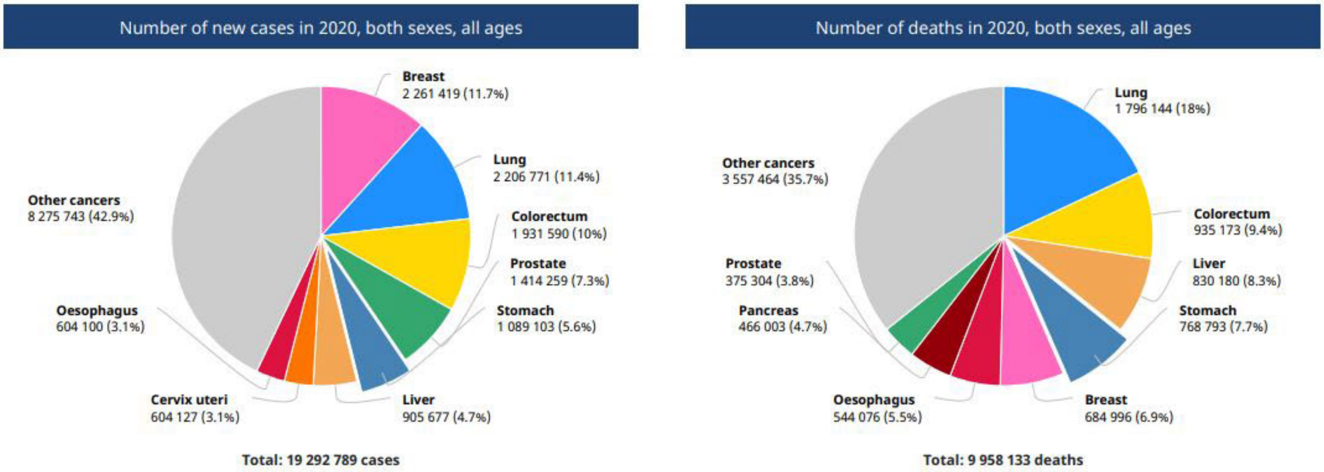

FIGURE 1. All cancer fact sheets - WHO; Data source - GLOBOCAN 2020 [11].

laborious and time-consuming task, and the output generated should be adapted to the perceptual quality of physicians so that diseases can be diagnosed promptly.

ML and DL models, as well as the datasets utilized in GI categorization and detection, are examined in this work. The research problems addressed in this paper include the following:

1) Q_1: Which ML models perform most effectively? *ML models are compared based on their performance. The performance metrics used in this work are also listed.*
2) Q_2: Which DL models perform most effectively? *DL models are compared based on their performance. The performance metrics used in this work are also listed.*
3) Q_3: What is the most commonly used modality for GI classification? *The most common image modality contributing to GI classification and detection is observed.*
4) Q_4: What datasets are available for GI classification? *All public datasets for GI endoscopy are listed and discussed.*

Q_1 is addressed in Section IV-B, Q_2 is addressed in Section IV-C, while Q_3 is observed in both Section IV-B and Section IV-C. Section IV-E is devoted to answer Q_4.

III. METHOD

Since their introduction to the medical community in the late 1970s, systematic reviews have become more prevalent. Systematic reviews have a special place in the medical community and are a useful tool for both academic research and clinical practice. They reveal information on knowledge gaps, which helps form the foundation for establishing practice standards and also drives the direction of future research projects. Well-written systematic reviews provide an efficient approach to assessing massive amounts of information. Before the review process starts, it is important to decide on the premise for conducting a systematic review and the methodology that will be employed.

A solid systematic review attempts to reduce bias and random error by summarizing the results of various primary research papers on a subject [14], [15]. A systematic review must meet three principles: explicit, rigorous, and replicable. These principles have endured the test of time and can be found in previous works that describe systematic review techniques. As a result, it is crucial to adhere to the already accepted published technique. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses Statement, or PRISMA Statement, is the most widely used reporting guideline for systematic reviews that address the literature review and search component [16], [17]. As a result, we decided to conduct this systematic review in accordance with the PRISMA guideline [15], [16], [17].

Typically, a systematic review starts with a specifically stated reflective research hypothesis or question. The inclusion and exclusion criteria for studies can be established after the research question has been created. The next procedure is conducting the comprehensive search by following these steps:

1) Determine in advance which databases and specific terms will be used to search in each database's keywords, title, and abstract fields.
2) Scour all suitable databases using the keywords and index terms from this literature search.
3) Investigate the references in the research found in the previous phases to broaden the search.

A. LITERATURE SEARCH STRATEGY AND IDEA VALIDATION

The literature search is essential to a systematic review [15]. The process of doing a literature search, also known as information retrieval, not only influences the findings of a systematic review but also establishes the data that may be

TABLE 1. **The significance of this review article compared to other review articles.**

Author(s)	Year	Objective	Limitations	Comparison with this review article
de Souza et al. [2]	2018	Review and investigate the feasibility and usage of ML techniques in the context of BE evaluation, dysplasia description, and treatment.	Focus on Barret esophagus and consider the works dating from 2011 to 2017.	Discuss ML and DL models for upper GI tract based on various WCE images and consider very recent works dating from 2015 to 2021.
Yang and Bang [3]	2019	Summarize AI applications in the field of GI.	Conduct a case-control study from a single center.	Study from well-designed multicenters.
Du et al. [4]	2019	Apply DL in GI images.	Limit to supervised learning for limited interest in Polyp, Hemorrhages, GI Cancer, MultiGI disease.	Discuss about supervised and unsupervised learning including transfer learning with wider scopes about the rest GI diseases that are not commonly studied such as HP infection and hookworm detection.
Rahim et al. [5]	2020	Conduct comprehensive reviews of the techniques for the detection of tumors, polyps, and ulcers.	Focus on lesion detection by reviewing DL articles by limiting the study to polyp/ulcer and hookworms.	Discuss ML and DL for various upper GI tract diseases found inside WCE images and detailed metrics in tabular forms.
Jha et al. [6]	2021	Present the overview of GI challenges held and imaging modalities used over past five years.	Limit to Medico 2017, Medico 2018 and BioMedia 2019 challenges.	Review and investigate the feasibility and usage of ML and DL techniques in the context of upper GI tract based on various WCE images from various sources.
Kim and Lim [7]	2021	Mainly discuss about binary classification and briefly about multiple lesion detection to compared DL performance with existing handcrafted ML performance.	Discuss specific diseases (GI Hemorrhage, Angioectasia, Erosion and ulcer (inflammation), Celiac disease, Polyp, tumor and Capsule localization.	Review and investigate the feasibility and usage of ML and DL techniques in the context of upper GI tract based on various WCE images from various sources.
Muruganantham and Balakrishnan [8]	2021	Conduct a systematic literature review on wireless capsule endoscopy image analysis that uses DL models.	Focus on anomaly classification, detection, and segmentation using supervised DL on WCE images.	Review and investigate the feasibility and usage of ML and DL techniques in the context of upper GI tract based on various WCE images from various sources.
Trasolini and Byrne [9]	2021	Review the major advances in AI for capsule endoscopy in recent publications.	1. Review publications from January 2000 to May 2020. 2. Focus on bleeding and small bowel detection; there is no discussion about the techniques of ML and DL used in the studies.	1. Review recent publication of AI for wireless capsule endoscopy (WCE) from January 2015 - March 2022 with wider scope of keywords. 2. Provide formal discussion section that can enlighten the insights of each technique of ML and DL.
Visaggi et al. [10]	2022	Report the performance of AI systems in the instrumental or clinical diagnosis of several esophageal and gastric diseases.	Discuss specific diseases, such as GI hemorrhage, angioectasia, erosion and ulcer (inflammation), celiac disease, polyp, tumor, and capsule localization.	Present usages of ML and DL techniques in the context of the upper GI tract based on various WCE images from various public datasets and wider scopes of keywords.

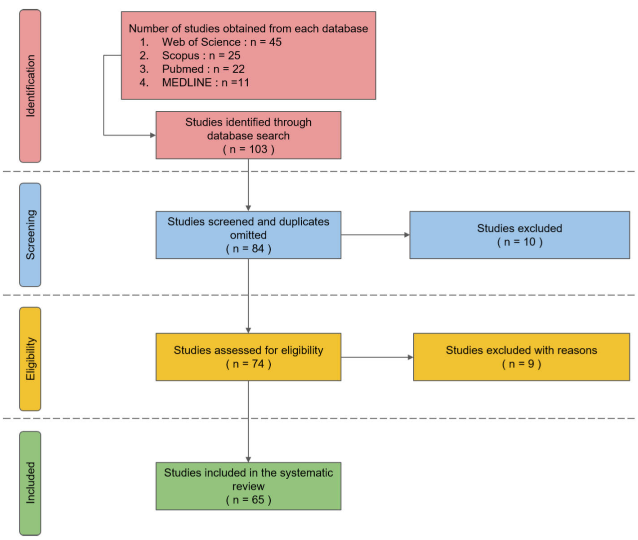

FIGURE 2. **The systematic review procedure by following PRISMA guideline [16], [17].**

used for analysis. This section presents how this process is conducted. By adhering to several practices outlined in [18], the search strategy and requirements are defined for this study's selection and exclusion. First, publications that have been published between January 2015 and March 2022 are taken into account. An automatic search is utilized in the four large databases available, Scopus, Web of Science, Pubmed, and MEDLINE, to retrieve relevant publications that have been published as journal articles or conference papers. The keywords extracted from several papers are the basis of the following search criteria.

(machine learning OR deep learning OR artificial intelligence OR machine intelligence OR computational intelligence OR image recognition OR convolutional neural network)
AND
(wireless capsule endoscopy OR video capsule endoscopy OR gastrointestinal diagnosis OR esophageal OR gastroduodenal endoscopy OR cancer OR carcinoma OR neoplasia OR Barrett's esophagus OR esophagitis OR gastro-esophageal OR ulcer detection OR bleeding detection OR motility disorder OR gastric cancer OR helicobacter pylori infection)

B. INCLUSION AND EXCLUSION CRITERIA

After the search phase, a number of publications are obtained that require additional screening. Table 2 shows the criteria for the inclusion and exclusion of all searched publications. Publications that meet the inclusion criteria should be published as journal articles, written in English, and discuss ML and DL techniques that utilize WCE images to detect lesions

TABLE 2. Set of inclusion and exclusion specifications for pertinent papers.

Type	Inclusion	Exclusion
Type of publication	Journal publications and conference proceedings	Ph.D. dissertations, working papers, editorial comments, technical reports
Language	English	Non-English
Scope	ML and DL for lesion detection in upper GI tract using WCE	Publications that fall outside the broader field of lesion detection in upper GI tract analysis, ML and DL not using WCE

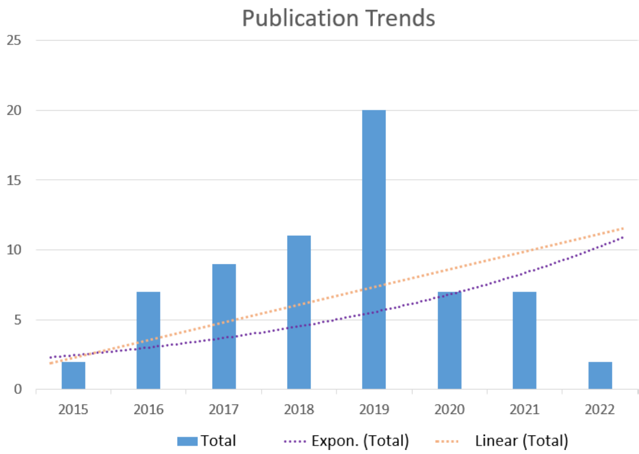

FIGURE 3. Publication growth from January 2015 to March 2022.

in the upper GI tract. In contrast, papers considered gray literature are not included in the final selection. In addition, papers written in languages other than English and papers that discuss the broader topic of lesion detection are excluded. Following the screening phase, 65 publications are finally considered for further analysis.

IV. RESULTS AND DISCUSSION

This section formally explains the methods used to classify WCE images for detecting various malignancies. The following section also includes a comparison of the methods that have been utilized for each malignancy classification.

A. RESEARCH TRENDS AND PUBLICATION VENUES

Figure 3 depicts the number of studies conducted throughout the period from January 2015 to March 2022 that is under consideration. Evidently, at least two studies on applying ML and DL algorithms to detect lesions in the upper GI tract have been conducted during this period. According to the trend, there has been an increased interest in employing ML and DL in the purview of lesion detection. The results reveal that there has been a significant surge in interest in ML and DL algorithms in 2019.

Table 3 summarizes the papers chosen for final inclusion (i.e., 65 studies) and the venues in which they are published. The table also recapitulates the distribution of the studies with respect to publication locations, publication types (either journals or conference proceedings), the number of corresponding studies, and relative fraction of total studies. Two recent journal performance metrics, CiteScore and Impact Factor from Scopus and Web of Science, respectively, are listed as well. Forty-four different publication sites exist in which the vast majority of papers have been published in the IEEE International Symposium on Biomedical Imaging (five papers). Other essential publication sites are the Annual International Conference of the IEEE Engineering in Medicine and Biology Society and Computers in Biology and Medicine, each of which published four papers. The top venues on this application domain in terms of Impact Factor metric are The Lancet Oncology, followed by The Lancet Digital Health, eBioMedicine, IEEE Transactions on Medical Imaging, and GI Endoscopy. Among these journals, however, only GI endoscopy published more publications (three papers) throughout the publication period (January 2015-March 2022).

B. SELECTED STUDIES THAT EMPLOY MACHINE LEARNING (ML) TECHNIQUES

In recent years, a proliferation of AI-based applications has rapidly transformed our work and life. AI can be defined as the creation and implementation of computer algorithms capable of performing activities that would otherwise necessarily involve the use of human intelligence. ML is a form of AI in which an algorithm uses input raw data to analyze features in a separate dataset and then produces a classified output as required. One of the most common uses of ML in medicine is image detection and classification. A training set of images containing the appropriate categories is used to train the system, resulting in better performance and fewer errors in traditional ML. The system's performance is tested using an independent set of images after a series of training sequences. Algorithms like support vector machine (SVM) and multi-layer perceptron (MLP) are common in traditional ML. Table 4 summarizes the findings of some studies that have used ML to help in GI diagnosis.

As presented in Table 4, SVM is widely used in GI image analysis. Sixteen papers apply SVM, while the other eight papers apply MLP, KNN, and random forest. According to investigated papers, the performance of SVM for detecting bleeding has a range of around 0.87 - 0.98 in terms of accuracy, 0.85 - 0.98 in terms of sensitivity, and 0.93 - 0.98 in

TABLE 3. Recapitulation of the selected studies according to their publication venues.

No	Publication venue	Type	#Studies	%	CiteScore (2021)	Impact factor (2021)
1.	IEEE International Conference on Robotics and Automation	Conference	1	1.54	-	-
2.	IEEE Transactions on Medical Imaging	Journal	2	3.08	18.8	11.037
3.	Healthcare Technology Letters	Journal	2	3.08	5.6	-
4.	IEEE International Conference on Real-time Computing and Robotics	Conference	1	1.54	-	-
5.	IEEE Signal Processing in Medicine and Biology Symposium	Conference	1	1.54	-	-
6.	IEEE Journal of Biomedical and Health Informatics	Journal	1	1.54	10.9	7.021
7.	International Conference on Internet of Things, Data, and Cloud Computing	Conference	1	1.54	-	-
8.	International Conference on Advanced Technologies for Signal and Image Processing	Conference	1	1.54	-	-
9.	International Workshop on Content-Based Multimedia Indexing	Conference	1	1.54	-	-
10.	IEEE Transactions on Automation Science and Engineering	Journal	1	1.54	11.0	6.636
11.	Biomedical Signal Processing and Control	Journal	3	4.61	6.9	5.076
12.	IEEE Journal of Translational Engineering in Health and Medicine	Journal	2	3.08	6.6	2.899
13.	Journal of Medical and Biological Engineering	Journal	1	1.54	3.5	2.213
14.	Telecommunications Forum	Conference	1	1.54	-	-
15.	Annual International Conference of the IEEE Engineering in Medicine and Biology Society	Conference	4	6.15	-	-
16.	Computers in Biology and Medicine	Journal	4	6.15	8.2	6.698
17.	International Conference on Advances in Science, Engineering, and Robotics Technology	Conference	2	3.08	-	-
18.	Conference of Open Innovations Association	Conference	1	1.54	-	-
19.	Cluster Computing	Journal	1	1.54	4.7	2.303
20.	Soft Computing	Journal	1	1.54	6.3	3.732
21.	The Lancet Digital Health	Journal	1	1.54	20.5	36.615
22.	Neurocomputing	Journal	1	1.54	10.3	5.779
23.	IEEE International Symposium on Biomedical Imaging	Conference	5	7.69	-	-
24.	eBioMedicine	Journal	1	1.54	15.0	11.205
25.	Scientific Report	Journal	1	1.54	6.9	4.996
26.	IEEE International Conference on Image Processing	Conference	2	3.08	-	-
27.	IEEE EMBS International Conference on Biomedical & Health Informatics	Conference	1	1.54	-	-
28.	IEEE International Conference on Machine Learning and Applications	Conference	1	1.54	-	-
29.	Sensors	Journal	1	1.54	6.4	3.847
30.	Gastrointestinal Endoscopy	Journal	4	6.15	9.8	10.396
31.	Journal of Medical System	Journal	1	1.54	11.5	4.920
32.	The Lancet Oncology	Journal	1	1.54	56.9	54.433
33.	Physics in Medicine & Biology	Journal	1	1.54	6.6	4.174
34.	Computational and Mathematical Methods in Medicine	Journal	2	3.08	2.8	2.809
35.	Multimedia Tools and Applications	Journal	1	1.54	5.3	2.577
36.	Journal of Medical Internet Research	Journal	1	1.54	8.2	7.093
37.	Pattern Recognition Letters	Journal	1	1.54	8.6	4.757
38.	Digestive and Liver Disease	Journal	1	1.54	5.6	5.165
39.	Computers, Materials & Continua	Journal	1	1.54	4.9	3.860
40.	Journal of Digital Imaging	Journal	1	1.54	7.4	4.903
41.	Complexity	Journal	1	1.54	3.5	2.121
42.	Clinical and Experimental Gastroenterology	Journal	1	1.54	5.0	-
43.	BMC Gastroenterology	Journal	1	1.54	3.2	2.847
44.	Clinical Endoscopy	Journal	1	1.54	3.4	2.05

terms of specificity. The results from the studies on KNN process indicate competitive performance with accuracy of 0.96 - 0.99, sensitivity of 0.92-0.99, and specificity of 0.96 - 0.99. All ML models use the standard RGB image modalities from private sources. Therefore, it is difficult to have an objective comparison for their performance. The major disadvantage of these conventional and handcrafted systems is the requirement to design and engineer a system for a specific task.

C. SELECTED STUDIES THAT EMPLOY DEEP LEARNING (DL) TECHNIQUES

DL and ML are instances of AI. Neural networks with many hidden layers are known as DL architectures. In recent years, the state-of-the-art performances of deep convolutional neural networks (DCNNs) have led to DL methods being recognized as the most sophisticated AI techniques. Image and video detection and classification are two domains where DL has been showing promising results and has become

TABLE 4. Classification of selected studies based on ML detection techniques and other relevant features between January 2015 and March 2020.

Author(s)	Year	Objective	Method	Image size	Angle	Dataset	Modality	Metrics
Yuan and Meng [19]	2015	Bleeding detection	Saliency map and SVM	256×256	140^0	Private	M4	Acc. : 0.95, Sen. : 0.98, Spec. : 0.93
Yuan et al. [20]	2015	Ulcer detection	Dense SIFT, dense HOG, and LLC	256×256	140^0	Private	M4	Acc. : 0.92, Sen. : 0.94, Spec. : 0.91
Charisis and Hadjileontiadis [21]	2016	Crohn disease	Curvelet transform, GA, and SVM	340×340	156^0	Private	M4	Sen. : 0.84 and Spec. : 0.83
Jia et al. [22]	2016	Bleeding detection	K-mean clustering, SVM, DBSCAN, and MapReduce Framework	240×240	-	Private	M4	R : 0.97, Prec. : 0.99, and F1 : 0.98
Maghsoudi et al. [23]	2016	normal, bleeding, diseases and tumors detection in the GI tract	Gabor, Haralick, Local binary pattern (LBP), Fisher test and MLP	26×26	-	Private	M4	Acc. : 0.97, Sen: 0.97, Spec.: 0.97, Prec : 0.95
Yuan et al. [24]	2016	Bleeding detection	SVM and KNN	180×180	-	Private	M4	Acc. : 0.96, Sen: 0.92, Spec.: 0.96, AUC : 0.98
Charfi and Ansari [25]	2017	Bleeding detection	SVM	256×256	140^0	Private	M4	Acc. : 0.94, Sen: 0.96, Spec. : 0.93
Charfi and El Ansari [26]	2017	Ulcer detection	SVM and MLP	243×424	-	Private	M4	Acc. : 0.94, Sen: 0.97, Spec. : 0.92
Suman et al. [27]	2017	Bleeding detection	SVM	-	-	Private	M4	Acc. : 0.98, Sen: 0.97, Spec. : 0.95
Yuan et al. [28]	2017	Bleeding, polyp, ulcer, normal detection	SVM	-	-	Private	M4	Acc.: 0.87
Deeba et al. [29]	2018	Bleeding detection	Region Growing, and SVM	-	-	Private	M4	Acc. : 0.95, Sen. : 0.94, Spec. : 0.95
Ghosh et al. [30]	2018	Bleeding detection	KNN	256×256	-	Private	M4	Acc. : 0.98, Sen : 0.99, Spec : 0.99
Ghosh et al. [31]	2018	Bleeding detection	SVM	-	-	Private	M4	Acc. : 0.98, Sen : 0.97, Spec: 0.98
Tuba et al. [32]	2018	Bleeding detection	SVM	-	-	Private	-	DSC : 0.85, ME : 0.09
Xing et al. [33]	2018	Bleeding detection	KNN	-	-	Private	M4	Acc : 0.99, Sen : 0.98, Spec : 0.99
Kundu and Fattah [34]	2019	Bleeding detection	KNN, NB, Decision Tree, Random Forest, SVM, ANN	5769×576	-	Private	-	Acc. : 0.97, Sen : 0.98, Spec : 0.96, Prec : 0.90
Mamun and Hossain [35]	2019	Ulcer detection	Logistic Regression Classifier	-	-	Private	-	Acc. : 0.88, Sen : 0.96, Spec : 0.76
Mamun et al. [36]	2019	Bleeding detection	WKNN	-	-	Private	-	Acc. : 0.98, Sen : 0.99, Spec : 0.99, Prec : 0.95, NPV : 0.99, F1 : 0.97
Obukhova et al. [37]	2019	Bleeding detection	LDA, SVM, RDF, AdaBoost	-	-	Private	-	Acc. : 0.95, Sen : 0.85, Spec : 0.97
Sivakumar and Kumar [38]	2019	Bleeding detection	Naïve Bayes	-	-	Private	M4	-
Ali et al. [39]	2020	detection of abnormalities in gastric	SVM	-	140^0	Private	-	Acc : 0.86, AUC : 0.91
Charfi and El Ansari [40]	2020	Ulcer detection	MLP, SVM and Random Forest	256×256	-	Private	-	Prec : 0.99, R : 0.99, F1 : 0.99
Kundu et al. [41]	2020	Multiple GI disease detection	M-LDA	-	-	Private	-	Acc: 0.91, Prec : 0.87, R : 0.85, F1 : 0.86
Rosenfeld et al. [42]	2020	GI detection and classification	logistic regression, decision tree, naive Bayes, SVM, random forest	-	-	Private	M6	Acc : 0.84, Sen : 0.90, Spec : 0.68, Prec : 0.84, R : 0.85, AUC : 0.87, F1 : 0.84

List of code. M1 = WLI, M2 = NBI, M3 = FL, M4 = Usual RGB, M5 = Hig-res, M6 = Non-image

increasingly popular. Because of the significant progress made in the image and video detection on large-scale annotated training sets, image and video recognition technologies are now being developed for use in various fields, including the medical field. Therefore, recent advances in medical image analysis have included DL techniques. These methods are typically classified into supervised learning and unsupervised learning. DL architectures usually employed in GI image analysis are trained with labeled data in a supervised setting. As presented in Table 5 and Table 6, almost all the literature related to deep neural networks (DNNs) used in GI image analysis is based on CNN (supervised learning), while only six papers apply other networks such as artificial neural network (ANN) and DNN.

Currently, detection by DL methods is a common task in GI image analysis. There are two object detection methods based on CNNs: single shot multibox detection (SSD) [43], [44], [45] and mask region-based CNN (Mask RCNN) [46], [47]. Both are popularly used in GI image analysis. The SSD method transforms object detection into an end-to-end target detection for regression problems. Mask RCNN combines region proposal algorithms and CNN classification. Additionally, generative adversarial network (GAN) [48], [49], which is an unsupervised architecture, also holds promise for GI image analysis. GAN is composed of two simultaneously trained and competing models: a generative model G that captures data distribution and a discriminative model D that estimates the probability that a sample comes from the training data rather than G. During the training procedure, G tries to maximize the probability of D making mistakes. This model is also described as a minimax two-player game. At the end, there is a unique solution: G recovers the training data distribution, and D equalizes to 0.5 everywhere. Both models, G and D, can be trained with backpropagation.

Training DNN from scratch requires large quantities of labeled data. The training and optimizing process of the network is usually very time-consuming. Collecting the required large number of GI images and having those images expertly annotated presents fundamental challenges, as label errors increase as experts become more fatigued. Hence, most GI image analysis tasks based on DL methods adopt a transfer learning approach [50], [51], [52], [53], which can reduce DNN's need for training data. In transfer learning, the model trained on a large image dataset, like ImageNet, is called the pre-trained model.

One transfer learning method is related to the feature extractor. The CNN layers of the pre-trained model are used as a feature extractor, and the fully connected layers of the pre-trained model replace the traditional classifier, like a linear classifier SVM. The GI image analysis tasks with small samples usually choose this transfer learning method. Another transfer learning method is the so-called fine-tuning. The input layer of the pre-trained model is replaced and trained by new data. One can fine-tune several or all layers of the pre-trained model. Typically, the previous DNN layers extract the images' generic features, such as edges and colors, which are useful for many tasks. The latter layers extract features related to a particular task, so the fine-tuning method often only fine-tunes the latter layers. The other transfer learning method is parameter sharing. The parameters of pre-trained DNN are loaded as the initialization parameters and trained with new data again, which can speed up the training process. The newly trained model shares the same network and parameters as the pre-trained model. This transfer learning method usually requires large training datasets.

D. SELECTED STUDIES BASED ON TYPES OF DISEASES

The stomach, which is located between the esophagus and the duodenum, is a vital component of the digestive system. The stomach is placed within the abdomen, between the cardiac and pyloric orifices of the GI system, and is covered by and connected to other organs by the peritoneum [86], [87]. Its principal role is to facilitate the mechanical and chemical assimilation of food that enters the stomach via the pharyngeal hole. In addition to processing, this organ is also involved in digestion and plays a vital role in regulating emission and motility in linked organs. Some of the most anatomically significant stomach sections are depicted in four major regions of the stomach: the cardia, fundus, body, and pyloric sections. The cardia covers the cardiovascular hole, which connects the throat to the stomach. The fundus, which is located in comparison to the flat plane of the cardiovascular entrance, is the most prominent expansion of the stomach. The most important part of the organ is the corpus or gastric body. The final segment of the stomach is the pyloric portion, which empties its contents into the duodenum. In addition, the pylorus is divided into two distinct zones: the pyloric antrum, which is connected to the stomach, and the pyloric channel, which is connected to the duodenum [87]. The stomach is primarily J-shaped, with two additional bends. The more elongated and elevated bend on the left side of the stomach is the cardiovascular indent, which is framed between the esophageal border and the fundus. The more limited inward bend contains a small notch called the angular incisures, which marks the line of intersection between the pyloric portion of the stomach and the rest of the body. Indications for the use of wireless capsule endoscopy have evolved over the years, with some becoming clear and well-established over time, while others have recently come to light and require additional verification.

1) GI bleeding.
2) Crohn's disease.
3) Small bowel tumor.
4) Surveillance of polyps.
5) Evaluation of abnormal small bowel imaging.
6) Barrett's esophagus (BE) and early esophageal adenocarcinoma (EAC)
7) Celiac disease.
8) Detection of Helicobacter pylori (Hp) infection and hookworm.
9) Gastric intestinal metaplasia (GIM).

Wireless capsule endoscopy is user-friendly for viewing the small intestine in high resolutions. There are a few preliminary findings on the use of video capsule endoscopy to detect celiac disease. The major and most evaluated indication for capsule endoscopy is GI bleeding. Patients with cryptic bleeding, probable small intestinal tumors, polyposis syndromes, Crohn's disease, hookworm, severe celiac disease, ulcer, and Helicobacter pylori illness, are currently evaluated by capsule endoscopy [88]. With an 8-fold magnification lens optical system, video capsule endoscopy may achieve a magnification similar to dissection microscopy, allowing it to examine the small intestinal villous structure in detail. In accordance with GI anatomy and the diagnostic purposes of WCE [88], [89], this subsection discusses the application of ML and DL to four major types of GI malignancies or diseases [90], as shown in Figure 4.

1) AI APPLICATION FOR GI BLEEDING

Any internal bleeding within the GI tract must be identified quickly. On the other hand, WCE generates many images, each of which has a varied brightness quality dependent upon position. This variation makes diagnosis more challenging for clinicians. Because bleeding within the GI tract can result in other potentially life-threatening conditions, including cancers, polyps, and ulcers, it is necessary to develop methods to diagnose bleeding within WCE efficiently.

To address this challenge, several methods have been proposed. Few methods involving a saliency map to extract color and textural information have been offered to locate bleeding within a frame of WCE images [19], [28]. The saliency map is derived from the physicians' perspective, and bleeding patterns emerge as red hues that differ from the normal colors perspective in the second stage. The saliency extraction method divides images into saliency areas and non-saliency regions; hence, these six features are individually extracted from saliency and non-saliency regions and concatenated to represent the WCE image's information. Following the representation of WCE images as twelve features. This approach yields promising results, with SVM achieving a classification performance of 0.96, 0.98, and 0.93 for accuracy, sensitivity, and specificity, respectively [19].

On the other hand, Charfi and Ansari [25] presents a simpler method by using a color-based segmentation in the HSV color space and utilizing color histograms, local binary patterns, and SVM. They proved with fewer or limited features; their techniques obtained promising results, with sensitivity, specificity, and accuracy of 0.96, 0.93, and 0.94, respectively. Another idea for limited features is proposed by Suman et al. [27]. It is worth noting that their approach certainly exhibits a highly promising performance in detecting bleeding within WCE images. By employing automated removal processing of dark and light blocks followed by image enhancement and subsequently utilizing SVM to detect bleeding, the method yields an impressive accuracy score of 0.97, a specificity of 0.95, and a sensitivity of 0.97.

In an effort to further enhance the accuracy of detecting bleeding areas in WCE frames, Deeba et al. [29] proposed the use of two improved SVM classifiers based on the RGB and

TABLE 5. Classification of selected studies based on DL detection techniques and other relevant features between August 2016 and October 2019.

Author(s)	Year	Objective	Method	Image size	Angle	Dataset	Modality	Metrics
Jia and Meng [54]	2016	Bleeding detection	CNN and SVM	240×240	-	Private	M4	R: 0.992, Prec. : 0.99, and F1 : 0.99
Segui et al. [55]	2016	motility events classification in the small intestine	VGG	128×128	-	Private	M4	Acc. : 0.96
Chen et al. [56]	2017	Organ classification : entrance, stomach, small intestine and colon	HMM and O-CNN (based on AlexNet)	256×256	-	Private	M4	Acc. : 0.89, Sen : 0.94, Spec. : 0.90
Jia and Meng [57]	2017	Bleeding detection	CNN + Handcrafted feature (k-mean clustering)	240×240	-	Private	M4	R : 0.91, Prec. : 0.94, F1 : 0.92
Jia and Meng [58]	2017	Bleeding segmentation	SVM and FCN	240×240	-	Private	M4	Acc. : 0.871, IoU : 0.77
Shichijo et al. [51]	2017	Helicobacter pylori detection (Gastric cancer)	GoogLeNet and transfer learning	244×244	-	Private	M4	Acc. : 0.87, Sen : 0.89, Spec. : 0.87
Sekuboyina et al. [59]	2017	Lesion detection	CNN	320×320	-	Private	M4	Sen : 0.71, Spec : 0.72, AUC : 0.79
Van Riel et al. [53]	2018	Esophageal Adenocarcinoma (EAC) detection	Transfer Learning with Convolutional Neural Networks	1600×1200	-	Public	M4	AUC: 0.92
Takiyama et al. [60]	2018	Anatomical Classification	GoogLeNet (CNN)	244×244	-	Private	M1	Sen : 0.98, Spec : 0.72, AUC : 1
Gadermayr et al. [61]	2018	Celiac Disease	CNN	128×128	-	Private	M4	No metrics or quantitative analysis at all
Ghosh et al. [62]	2018	Bleeding detection	SegNet	360×360	-	Private	M4	Acc : 0.94, IoU : 0.90
Iakovidis et al. [63]	2018	Anomaly Detection	Weakly-supervised CNN (WCNN)	489×409 or 360×360	-	Public	M4	Acc. : 0.90, Sen : 0.93, Spec : 0.87
Pogorelov et al. [49]	2018	Angiectasia detection	V-GAN	520×520	156^0	Private	M4	Acc. : 0.99, Sen : 0.88, Spec : 0.99
Shvets et al. [64]	2018	Angiodysplasia detection	U-Net,TernausNet-11,TernausNet-16, and AlbuNet-34	576×576	156^0	Private	M4	DSC : 0.85, IoU : 0.75
Alaskar et al. [65]	2019	Ulcer detection	GoogLeNet and AlexNet	256×256	-	Private	M4	Acc. : 1, Sen : 1, Spec : 1, AUC : 1
Cai et al. [66]	2019	esophageal squamous cell carcinoma (ESCC) localization	DNN	-	-	Private	M1	Acc : 0.91, Sen : 0.98, Spec : 0.85, PPV : 0.86, NPV : 0.97
Cogan et al. [67]	2019	GI tract classification	Inception-ResNet-v2, Inception-v4, and NASNet	299×299	-	Public	-	Scc : 0.98, Prec : 0.94 , R : 0.94, Spec : 0.99, F1 : 0.94, MCC : 0.93
Hajabdollahi et al. [68]	2019	Bleeding Segmentation	CNN	-	-	Private	-	DSC : 0.89, AUC : 0.98
Lee et al. [69]	2019	Ulcer classification	Y-net (two VGG19)	-	-	Private	-	Acc : 0.96, Prec : 0.95, R : 0.93
Horie et al. [43]	2019	esophageal cancer detection	DNN (Single Shot MultiBox Detector)	300×300	-	Private	M1 and M2	Acc. : 0.95, Sen : 0.98, Spec : 0.79, PPV : 0.39, NPV : 0.95
Hajabdollahi et al. [70]	2019	Bleeding detection	CNN	-	-	Private	-	Acc : 0.99, Sen : 0.95, Spec : 0.99, DSC : 0.89, AUC : 0.99

List of code. M1 = WLI, M2 = NBI, M3 = FL, M4 = Usual RGB, M5 = Hig-res

TABLE 6. Classification of selected studies based on DL detection techniques and other relevant features between November 2019 and March 2022.

Author(s)	Year	Objective	Method	Image size	Angle	Dataset	Modality	Metrics
Khan et al. [71]	2019	stomach abnormalities detection and classification	DenseNet,Deep CNN	760×1240	-	Private	-	Acc : 0.99, R : 0.99, Spec : 0.99, Prec : 0.99, AUC : 0.99, FPR : 0.003
Klang et al. [72]	2019	Crohn disease ulcers	Xception CNN	516×516	-	Private	-	Acc : 0.96, Sen : 0.97, Spec : 0.96, PPV : 0.94, NPV : 0.98, AUC : 0.99
Luo et al. [73]	2019	upper GI cancer detection	GRAIDS algorithm (based on DeepLab's V3+)	-	-	Private	M5	Acc. : 0.93, Sen : 0.94, Spec : 0.92, PPV : 0.81, NPV : 0.98
Wang et al. [74]	2019	Ulcer detection	RetinaNet	480×480	-	Private	M4	Acc. : 0.90, Sen : 0.89, Spec : 0.90, AUC : 0.95,
Wang et al. [75]	2019	Ulcer detection	Hanet (based on ResNet-34) ;DeepCNN	480×480	-	Private	-	Acc. : 0.92, Sen : 0.92, Spec : 0.92, Prec : 0.90, R : 0.92, F1 : 0.92, AUC : 0.97
Xing et al. [76]	2019	upper GI cancer detection	SHNet (based on DenseNet)	576×576	-	Private	-	Prec : 0.95, R : 0.97, F1 : 0.96
Xu et al. [44]	2019	GI detection and classification	SSD network	300×300	-	Private	M2 and M4	Acc : 0.98, F1 : 0.99
Yang et al. [77]	2019	GI lesion detection	DCNN	256×240	-	Private	M4	mAP: 0.99
Khan et al. [47]	2020	GI diseases classification (ulcer, bleeding, polyp, and healthy)	Modified Mask-RCNN	-	-	Private	-	Acc. : 0.99, Prec : 1 , IoU : 0.88
Ozturk and Ozkaya [78]	2020	GI classification	CNN+ANN and CNN+SVM	299×299	-	Public	-	Acc. : 0.97, Sen : 0.92, Spec : 0.99, Prec : 0.94, F1 : 0.93
Zhang et al. [50]	2020	chronic atrophic gastritis detection	CNN-CAG (based on DenseNet) and transfer learning	512×512	-	Private	M4	Acc : 0.94, Sen : 0.95, Spec : 0.94
Attique Khan et al. [52]	2021	stomach classification	ACO, deep transfer learning-based features, MLNN	512×512	-	Public	M4	Acc. : 0.99, Sen : 0.99, Prec : 0.96, F1 : 0.98
Bang et al. [79]	2021	Gastric neoplasm classification	AutoDL	640×480	-	Private	M1	Acc. : 0.87, Prec : 0.84, R : 0.68, F1 : 0.75
Ghosh and Chakareski [80]	2021	Bleeding detection	AlexNet and SegNet	576×576 and 256×256	-	Public	M3	Acc. : 0.99, Sen : 0.97, Spec : 0.99, Prec : 0.99, F1 : 0.98
Hmoud Al-Adhaileh et al. [81]	2021	GI disease detection	GoogleNet, ResNet-50, and AlexNet	720×576 and 1920×1072	-	Public	-	Acc : 0.97, Sen : 0.97, Spec : 0.99, AUC : 0.99
Klang et al. [82]	2021	gastric ulcer classification (benign or malign)	Xception CNN	299×299	-	Private	-	Sen : 0.92, Spec : 0.75, AUC : 0.91
Yogapriya et al. [83]	2021	GI tract disease classification	VGG16, ResNet-18, and GoogLeNet	720×576 and 1920×1072	-	Public	-	Acc : 0.96, R : 0.96, Prec : 0.96, F1 : 0.97, MCC : 0.95, KAP : 0.96
Xu et al. [84]	2021	GIM detection	VGG16, ResNet-50, DenseNet-169, and EfficientNet-B4	224×224	-	Private	M1 and M2	Acc : 0.91, Sen : 0.88, Spec : 0.93, PPV : 0.93, NPV : 0.89, AUC : 0.96
Siripoppohn et al. [85]	2022	GIM segmentation	BiSeNet	1350×1080	-	Public and Private	M1 and M2	Sen : 0.93, Spec : 0.80, PPV : 0.82, NPV : 0.92, Acc : 0.87, and IoU : 0.57
Siripoppohn et al. [85]	2022	GIM segmentation	U-Net	-	-	Private	M1 and M2	Sen : 0.88, Spec : 0.93, PPV : 0.90, NPV : 0.92, Acc : 0.91, AUC : 0.95, and KAP : 0.82

List of code. M1 = WLI, M2 = NBI, M3 = FL, M4 = Usual RGB, M5 = Hig-res

HSV color spaces. The classifiers are based on the RGB and HSV color spaces. The image regions are defined using statistical features taken from the first-order histogram probability of the corresponding color channels. Although their method exhibits an average accuracy of 0.95, a sensitivity of 0.94, and a specificity of 0.95, it falls slightly short compared to

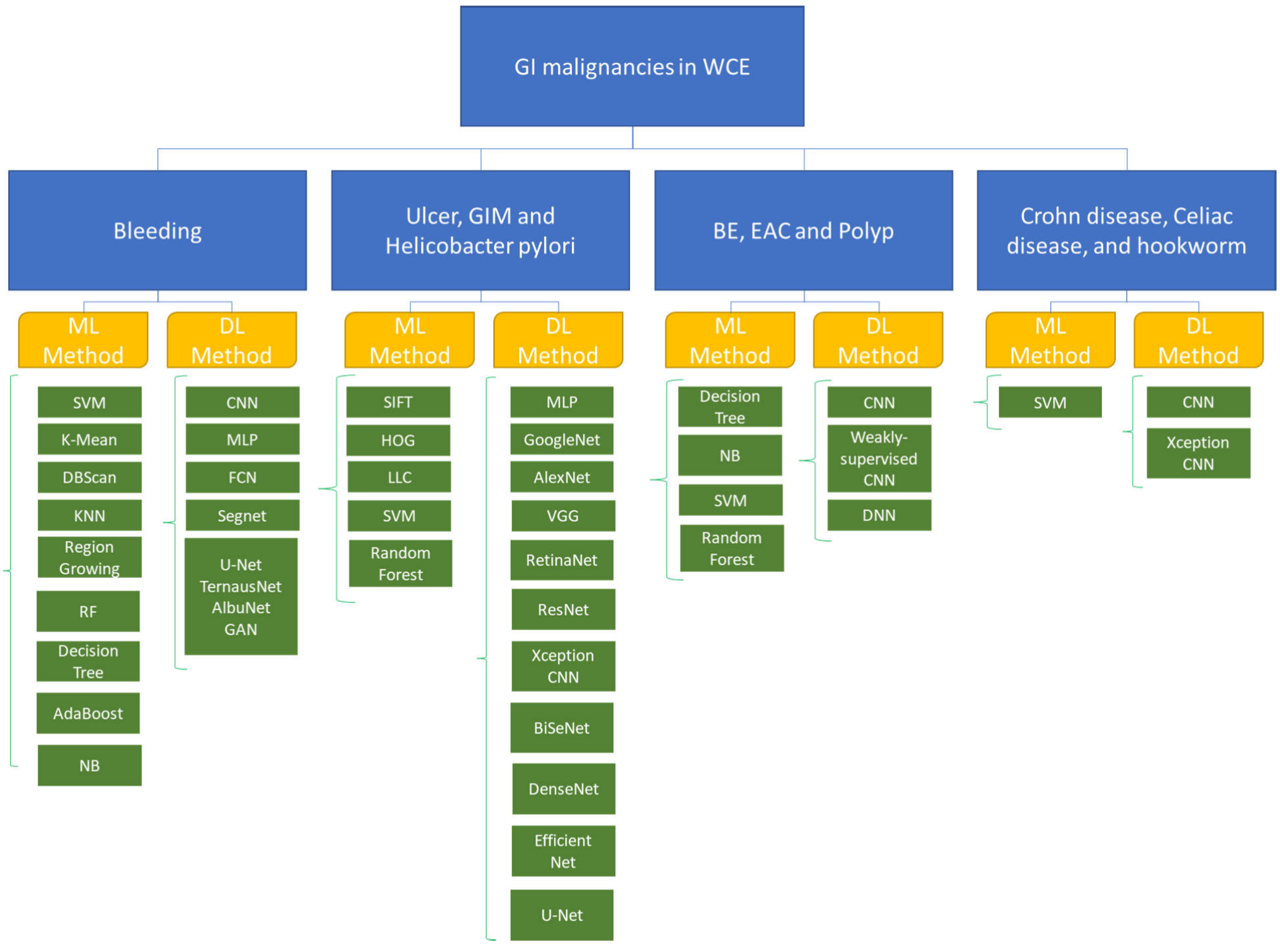

FIGURE 4. GI malignancy detection in WCE.

the performance of Suman et al. [27]'s method. Nonetheless, these studies offer valuable insights into the development of efficient methods for detecting bleeding within WCE images with limited features.

Another interesting idea for GI bleeding detection is proposed by Yuan et al. [24]. They offer a different approach by using an extension of the Bag of Words method and define WCE images as word-based color histograms. They use the word-based color histograms and K-means clustering approach on pixel images to generate cluster centers. Next, they use SVM and KNN algorithms to determine the state of a WCE frame. In contrast, Ghosh et al. [30] proposes using color histograms of block statistics (CHOBS) by utilizing block-based local feature extraction from each color RGB plane, leading to superior feature representation. The extracted features' dimensionality is reduced by using two stages of feature reduction: histogram pattern and PCA. KNN classifier is then used to determine the state of a WCE frame.

In summary, various methods in ML have been proposed to locate bleeding within Wireless Capsule Endoscopy (WCE) images. These methods utilize color and textural information extracted from multiple color spaces, including RGB, CIELAB, CIEXYZ, YUV, YIQ, CMYK, HSV, and HSI [19], [24], [25], [30]. Selecting the right color space is essential for WCE picture salient region extraction. Bleeding frame detection is also affected by this setting. These methods also employ a variety of classifiers, including SVM, KNN, Fisher score tests, and neural networks, for bleeding classification regions [19], [24], [25], [27], [28], [30], [54]. Additionally, Gabor filters, local binary patterns, haralick, color histograms, and feature reduction methods, such as PCA, are used to enhance the detection of bleeding regions [25], [30], [54]. Notably, the RGB color space has been found to be the optimal color space for bleeding detection in GI diagnosis, and the combination of SVM, KNN, and PCA achieves the best results. The best result of the proposed methods achieves 0.98 accuracy, 0.99 sensitivity, 0.99 specificity in the case of bleeding frame detection, and 0.96 precision in the case of bleeding zone detection [24], [30]. Furthermore, it has been found that ML models with fewer features perform better than models with more complex features [19], [25], [27], [29].

One significant breakthrough in the field is the introduction of a computerized bleeding detection approach at the pixel level, combined with the MapReduce framework [22]. This method comprises two stages: bleeding frame identification and region segmentation. K-means clustering algorithm and SVM are used to retrieve accurate bleeding frames, and DBSCAN is utilized to identify bleeding locations quickly. Finally, the MapReduce framework rewrites the entire workflow in distributed computing to speed up the discrimination process without sacrificing detection accuracy. This method achieves an accuracy of 0.97, a precision of 0.99, and an F1 score of 0.98. Another breakthrough is the use of transform color domains instead of RGB for pixel-based holistic feature extraction to improve bleeding frame detection in WCE [31]. Higher and lower-order statistical analyses are carried out on composite color domains to derive WCE picture features, resulting in very competitive results with an accuracy of 0.97, sensitivity of 0.97, and specificity of 0.98. Their research open-up new directions to improve GI bleeding detection using the transform color method instead of using raw RGB information.

Recently, with the rapid development of DL, researchers proposed several models based on CNNs to distinguish bleeding and non-bleeding WCE images through V-GAN [49], semantic segmentation method known as SegNet [62], U-Net, TernausNet, and AlbuNet34 [64] as well as transfer learning with pre-trained AlexNet [80]. Surprisingly, the optimal color space for these models has been shown to be Hue, saturation, and value (HSV) instead of RGB. However, the results demonstrate that the performance of DL models is inferior to that of the ML models for GI bleeding detection. Their results show a global accuracy in the range of 0.83-0.94 and an intersection over union (IoU) in the range of 0.75-0.91.

Therefore, to overcome this limitation, researchers devised an idea for a new framework by combining ML and DL models together [57], [58], [68], [70]. Jia and Meng [57], [58] describes a new approach to WCE bleeding detection that incorporates both handcrafted (HC) and CNN features. The framework has several stages: feature extraction, feature integration, and classification. In the feature extraction stage, the input frame is processed for both CNN-derived and HC features. Next, both results from the feature extraction stage are merged with a specified technique. The classification stage uses the combined feature vector as an input. A softmax classifier is used to make the final decision. Their proposed method of integrating handcrafted and CNN features achieves remarkable improvements compared to DL models alone, with an accuracy range of 0.95-0.99, an AUC score above 0.97, an IoU score of 0.98, and an F1 score of 0.93. This breakthrough provides new insights for the future development of GI bleeding detection [57], [58], [68], [70].

2) AI APPLICATION FOR ULCER, GASTRIC INTESTINAL METAPLASIA (GIM) AND HELICOBACTER PYLORI (Hp)

Ulcers are a prevalent medical condition that affects a large number of people around the world, and early detection and treatment can substantially improve the outcome for patients. AI have the potential to enhance the accuracy and efficiency of ulcer detection, which could result in earlier diagnosis and better treatment outcomes for patients. Several ML methods have been proposed to detect ulcers from WCE images by utilizing saliency map, Locality-constrained Linear Coding (LLC), K-means, Complete Local Binary Pattern (CLBP), and Global Local Oriented Edge Magnitude Pattern (Global LOEMP). Their best performance achieves an accuracy score in the range of 0.80-0.94 and specificity in the range of 0.91-0.94 in the experiments [20], [26], [28].

In recent years, there has been rapid development in deep learning, and researchers have been using these methods to develop new ulcer detection models. The use of deep learning models for ulcer detection has yielded remarkable results, with various studies showcasing the superior performance of different models such as AlexNet, GoogLeNet, VGG19, RetinaNet, ResNet, Mask-RCNN, and Xception CNN [65], [69], [74], [75], [82]. These models have demonstrated high accuracy, sensitivity, and specificity in detecting and classifying ulcers in medical images.

Among the methods discussed, Alaskar et al. [65] achieved the highest performance with an accuracy, sensitivity, and specificity of 1.00 using DL networks such as AlexNet and GoogLeNet. However, it is important to note that a newer model proposed by Khan et al. [47] that is based on Mask-RCNN achieved a classification accuracy of 0.99 and a mean overlap coefficient score of 0.88, which is also a highly advantageous result. Furthermore, new research has advocated using GoogLeNet and transfer learning to detect Helicobacter pylori infection. As proposed by Shichijo et al. [51], the impressive attained accuracy, sensitivity, and specificity of 0.83 demonstrate the potential of this method to identify the presence of this infection accurately.

Gastric cancer is a significant global health problem, ranking as the fifth leading cause of cancer-related death worldwide. The development of gastric adenocarcinoma (GAC) is often preceded by intestinal metaplasia (IM) in the cardia gastric mucosa. Therefore, detecting IM early on is crucial for preventing the development of (GAC). Recent studies have demonstrated the potential of using deep learning (DL) models for detecting gastric intestinal metaplasia (GIM) and chronic atrophic gastritis (CAG), which are indicators of gastric cancer.

One promising study, conducted by Xu et al. [84], proposed using several DL models, including ResNet-50, VGG-16, DenseNet-169, and EfficientNet-B4, to detect GIM. The authors achieved impressive results with their system, which was trained on multicenter data from five hospitals in China, achieving an accuracy of 0.91, sensitivity of 0.88, and specificity of 0.93. Their research highlights the potential of DL models to accurately detect early-stage gastrointestinal (GI) tract diseases, such as GIM. Another breakthrough idea is a real-time GIM segmentation method that uses a bilateral segmentation network (BiSeNet), proposed by Siripoppohn et al. [85]. Their method incorporated contrast-limited adaptive histogram equalization to boost the contrast of GIM regions. Their method achieved impressive sensitivity, specificity, PPV, NPV, accuracy, and IoU scores of 0.93, 0.80, 0.82, 0.92, 0.87, and 0.57, respectively.

In addition to GIM detection, recent research has also focused on developing DL-based video monitoring systems to diagnose chronic atrophic gastritis (CAG), another significant indicator of gastric cancer. A recent study by Zhao and Chi [91] developed a real-time DL-based video monitoring system that outperformed an experienced group of endoscopists, achieving a sensitivity of 0.88, specificity of 0.93, PPV of 0.91, NPV of 0.92, the accuracy of 0.91, AUC of 0.92, and KAP of 0.82. These impressive results highlight the potential of DL-based video monitoring systems for diagnosing early-stage gastric cancer.

In conclusion, despite the considerable expertise of endoscopists, relying solely on their visual diagnosis can pose significant challenges in the diagnosis and differential diagnosis of early malignant tumors and cancer detection. According to studies by Siripoppohn et al. [85] and Zhao and Chi [91], up to 10% of malignant lesions can be missed with conventional endoscopic methods. However, computer-aided diagnosis (CAD) using DL models has demonstrated significant potential for early detection of gastrointestinal tract diseases, including GIM and CAG, leading to improved gastric cancer diagnosis and prevention. Recent studies have proposed breakthrough ideas such as transfer learning and real-time segmentation, which demonstrate rapid advancements in this field. These advancements are likely to significantly impact the field of gastroenterology and contribute to reducing the global burden of gastric cancer.

3) AI APPLICATIONS FOR BARRETT'S ESOPHAGUS (BE), ESOPHAGUS CANCER, AND POLYPS

Esophageal cancer is one of the deadliest forms of cancer and is the sixth leading cause of cancer-related deaths globally. The disease progresses through various phases and becomes more challenging to treat as it advances. Barrett's esophagus is a precancerous condition that may develop into esophageal adenocarcinoma (EAC). Early detection of EAC is crucial, as endoscopic removal of malignant tissue can lead to five-year survival rates exceeding 95%.

The detection of early esophageal cancer using ML and DL methods has become an important area of research due to the high mortality rate associated with this disease. In the case of Barrett's esophagus, ML methods such as logistic regression, decision tree, naive Bayes, SVM, and random forest have been used to identify this precancerous condition, with performance scores ranging from 0.84 to 0.85 [42].

Meanwhile, the DL models have shown impressive improvement in the detection of Early Esophageal Adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC). The application of transfer learning [53], Single Shot MultiBox Detector architecture [43], and GI AI Diagnostic System (GRAIDS) [73] were breakthrough ideas to detect signs of early esophageal cancer that resulted in high-performance scores. These models achieved high sensitivity scores in the range of 0.94-0.98, accuracy scores in the range of 0.91-0.98, and specificity scores in the range of 0.85-0.92 [43], [53], [66], [73].

The results of the studies discussed above highlight the potential of both ML and DL approaches for the early detection of esophageal cancer. It is noteworthy that DL models consistently outperformed the ML models, suggesting that DL techniques may have an advantage and be particularly effective in this domain. These findings underscore the promise of AI, particularly in the field of gastroenterology, and contribute to reducing the global burden of esophageal cancer. Moving forward, it also provides a foundation for the continuation of future research and development in this area which is critical to realizing the full potential of AI for the early detection of esophageal cancer.

4) AI APPLICATIONS FOR CROHN'S DISEASE, CELIAC DISEASE, AND HOOKWORM

Several less-studied GI diseases, including Crohn's disease, celiac disease, and hookworm, pose a significant challenge in terms of accurate detection. To address this, Charisis and Hadjileontiadis [21] introduce a novel approach that utilizes the spatial morphology distribution domain, in conjunction with the space-frequency of curvature structures, to detect Crohn's disease. They employ the hybrid adaptive filtering (HAF) technique to recover lesion-related structural properties and reconstruct more meaningful images in the space-frequency domain. The differential lacunarity analysis is used to establish the spatial morphological distribution domain and extract textural features. SVM is utilized for the classification stage, with the YCbCr space's Cr channel being used as the input. The proposed approach achieves a sensitivity of 0.77 and a specificity of 0.86, demonstrating the potential of this method for detecting Crohn's disease.

In another study, Maghsoudi et al. [23] proposes a method for identifying Crohn's disease that utilizes textural characteristics, such as Gabor filters, local binary patterns, and Haralick in HSV color spaces, combined with Fisher score tests and neural networks. Their approach achieves an impressive accuracy, sensitivity, and specificity of 0.95, 0.97, and 0.94, respectively. These studies have shown that ML methods are particularly effective in this domain.

In a different study, Gadermayr et al. [61] proposes a DL method by utilizing patch-based CNNs as a replacement for traditional hand-crafted machine learning methods in the automated diagnosis of celiac disease. Their CNN is pre-trained on ImageNet, and their best performance achieves an accuracy of 0.90. Overall, these studies offer novel insights into the detection of less-studied GI diseases, particularly Crohn's disease and celiac disease. The utilization of advanced computer vision techniques such as spatial morphology distribution domain, space-frequency of curvature structures, traditional machine learning algorithms such as SVM, and deep learning such as patch-based CNNs, has shown promising results in the detection of these diseases. Future research could explore the integration of these

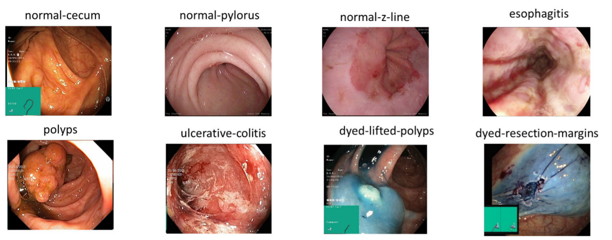

FIGURE 5. The KVASIR dataset's sample images of eight different classes.

techniques with other emerging technologies and image segmentation algorithms to further improve the accuracy and reliability of these less-studied GI disease detection within the GI tract.

E. PUBLIC DATASETS

This section explains various publicly available image and video datasets of the human GI tract. These public datasets are commonly employed in CAD systems that are used for GI examination.

1) KVASIR DATASET

There are four hospitals in Norway that participated in the medico challenge with a dataset of 8,000 endoscopic images, which have been annotated and validated by endoscopists from each of the four institutions. It has been the first time that an endoscopic GI dataset with eight classes is used in the Medico 2017 competition. It is known as the ''KVASIR'' dataset, which is a multi-class dataset with 1,000 images per class [92]. This is the first comprehensive dataset that mimics various endoscopic procedures as a whole.

The KVASIR dataset consists of multiple endoscopic findings for the entire GI tract. This dataset is classified into eight different anomalies, where each class consists of 1,000 images. These classes are dyed-lifted-polyps, normal-cecum, normal-pylorus, normal-z-line, esophagitis, polyps, ulcerative colitis, and dyed-resection-margins (see Figure 5). The various diseases and their respective class label encodings are listed in Table 7.

Datasets for training and testing have been created separately for Medico 2017. The number of images in the training and testing sets is approximately 4,000. Pre-split train-test categories with 500 images per class for each split are given to the participants (see Figure 6). There are, however, no available labels for the test set. The resolution of the images, captured with an Olympus endoscope with high resolutions, ranges from 720 × 576 pixels to 1920 × 1072 pixels. A green box in the left-bottom corner of some photos in the collection indicates the location of the scope inside a patient's digestive tract.

2) MEDICO 2018 AND BIOMEDIA 2019 CHALLENGE DATASETS

The dataset for the Medico 2018 challenge includes the images from the previous competition in addition to 6,033 more images and eight new classifications. Colon-clear, stool-inclusions, stool-plenty, blurry-nothing, out-of-patient, dyed-lifted-polyps, dyed-resection-margins, and the instrument class are the extra classes utilized in the task. The total number of endoscopic pictures and annotations for Medico 2018 is 14,033 [93].

The BioMedia 2019 Challenge dataset consists of a combination of the Medico 2018 Challenge dataset [93] and six video datasets [94]. These challenges with their extended dataset of 14,033 GI endoscopy frames and six video datasets are aimed at classifying 16 class categories for multiple GI endoscopy organs [94]. The dataset sample can be seen in Figure 7. The video dataset in BioMedia 2019 has various lengths from 51 seconds to 5 minutes and 11 seconds (see the details in Table 8). In most high-resolution GI endoscopes, the standard frame rates are over 45 frames per second.

Overall, the dataset is highly uneven (see Figure 8), with polyps constituting the majority class (1,625 photos; 11.58%) and out-of-patient constituting the minority class (10 images; 0.08%). This unbalanced distribution can be viewed as part of the problem and reflects actual hospital data collection. The included images range in resolution from 720 × 576 to 1920 × 1072 pixels. The dataset is divided into a training dataset comprising 5,293 images and a test dataset containing 8,740 images.

3) HYPER-KVASIR DATASET

On the basis of the lessons acquired from the publication of the KVASIR dataset and the organization of challenges, it is evident that data availability remains one of the most significant obstacles in medical AI. It is difficult to retrieve data from healthcare systems; approval from medical committees is challenging to obtain, because medical specialists have limited time, and there is no efficient labeling tool for such data. Consequently, Borgli et al. [95] create Hyper-KVASIR.

The photos and videos contained in Hyper-KVASIR are collected prospectively from 2008 to 2016 from routine clinical examinations conducted in a Norwegian hospital (see Figure 9). The photos have been retrieved from the Picsara image documentation database (CSAM, Norway) in 2016, which is an add-on to the electronic medical record system. With Hyper-KVASIR, both the quantity of labeled medical data for supervised learning and unlabeled data increases significantly. The 1.17 million images and frames in the new dataset come from 110,079 images and 374 videos of different GI exams. The details of the Hyper-KVASIR dataset are presented in Table 9. All various 23 classes are shown in Figure 9, and the dataset class distribution is shown in Table 10.

4) KID DATASET

The KID dataset is a publicly available database of annotated WCE images and videos, including pixel-level annotations. It contains WCE images obtained from the whole GI tract using a Miro-Cam capsule endoscope with a resolution of 360×360 pixels. These include 303 images of

TABLE 7. Details of KVASIR disease dataset.

Disease	Class	Description
Normal-cecum	Class 0	In the lower abdominal cavity, the cecum is a long tube-like structure. It usually gets foods that have not been digested. The significance of identifying the cecum is that it serves as evidence of a thorough colonoscopy.
Normal-pylorus	Class 1	The pylorus binds the stomach to the duodenum, the first section of the small bowel. The pylorus must be located before the duodenum can be instrumented endoscopically, which is a complicated procedure.
Normal-Z-line	Class 2	The Z-line depicts the esophagogastric junction, which connects the esophagus's squamous mucosa to the stomach's columnar mucosa. It is vital to identify Z-line to determine whether or not a disease is available.
Esophagitis	Class 3	Esophagitis is a condition in which the esophagus becomes inflamed or irritated. They appear as a break in the mucosa of the esophagus.
Polyps	Class 4	Polyps are clumps of lesions that grow within the intestine. Although the majority of polyps are harmless, a few of them can lead to colon cancer. As a result, detecting polyps is essential.
Ulcerative colitis	Class 5	The entire bowel is affected by ulcerative colitis (UC), which can lead to long-term inflammations or bowel wounds.
Dyed lifted polyps	Class 6	The raising of the polyps decreases the risk of damage to the GI wall's deeper layers due to electrocautery. It is essential to pinpoint the areas, where polyps can be removed from the underlying tissue.
Dyed resection margins	Class 7	The resection margins are crucial for determining whether or not the polyp has been entirely removed.

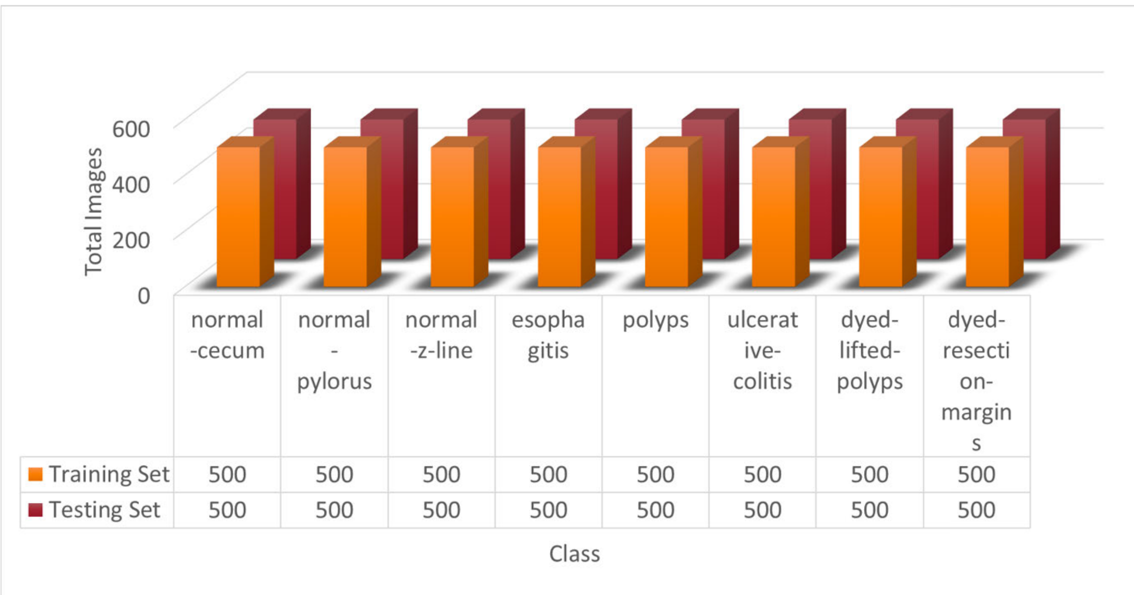

FIGURE 6. The KVASIR dataset class distribution for Medico 2017.

vascular anomalies (e.g., small bowel angiectasias, lymphangiectasias, and blood in the lumen); 44 images of polypoid anomalies (e.g., lymphoid nodular hyperplasia, lymphoma, Peutz-Jeghers polyps); 227 images of inflammatory anomalies (e.g., ulcers, aphthae, mucosal breaks with surrounding erythema, cobblestone mucosa, luminal stenoses and/or fibrotic strictures, and mucosal/villous oedema); and 1,778 normal images obtained from the esophagus, the stomach, the small bowel and the colon. This dataset total is 2,352 images.

5) KVASIR CAPSULE DATASET

The Kvasir Capsule dataset [96] is an extensive capsule endoscopy dataset compiled from hospital examinations in Norway from February 2016 to January 2018. The dataset has been collected using the Olympus Endocapsule 10 System

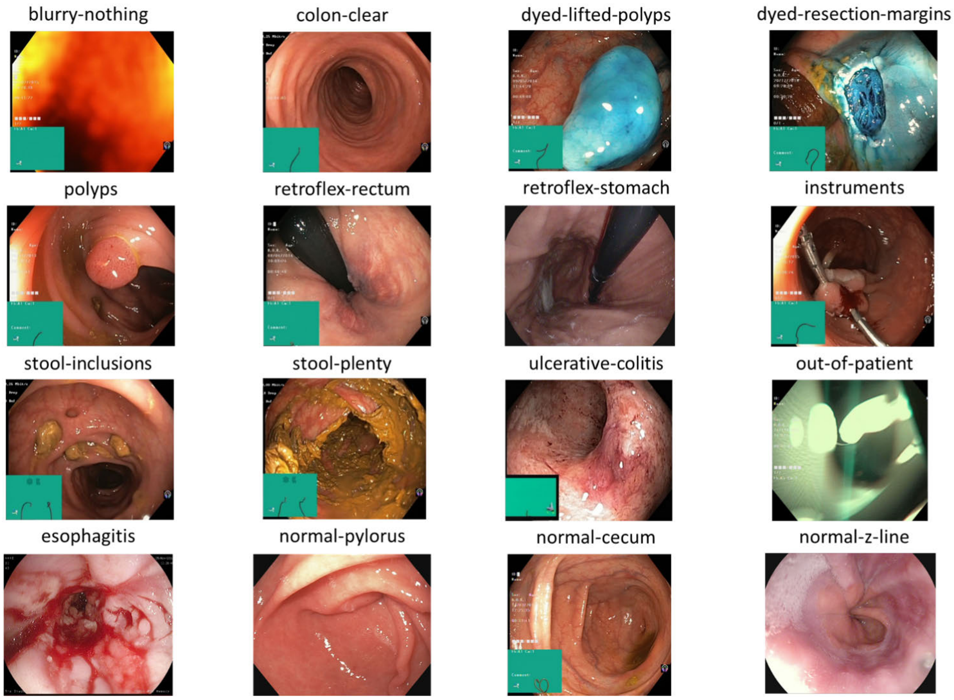

FIGURE 7. Medico 2018 and BioMedia 2019 dataset classes and sample images.

FIGURE 8. The BioMedia 2019 dataset class distribution.

with a frame rate of 2 and 30 frames per second. The Olympus Endocapsule system captures in total between 50 and 100 thousand frames, with pixel resolutions of 336 × 336. The Kvasir Capsule consists of 117 videos from which 4,741,504 images can be extracted. It contains 47,238 labeled and medically confirmed images by four medical specialists, with a bounding box surrounding findings from 14 distinct classes. In addition to these labeled images, the dataset contains 4,694,266 unlabeled images. The various diseases and their respective class label encodings are listed in Table 10.

All 14 various labeled classes are shown in Figure 11, and the details of the KVASIR Capsule dataset are shown in Table 11. Additionally, as shown in Figure 12, the quantity of images per class is unbalanced, which is a common issue

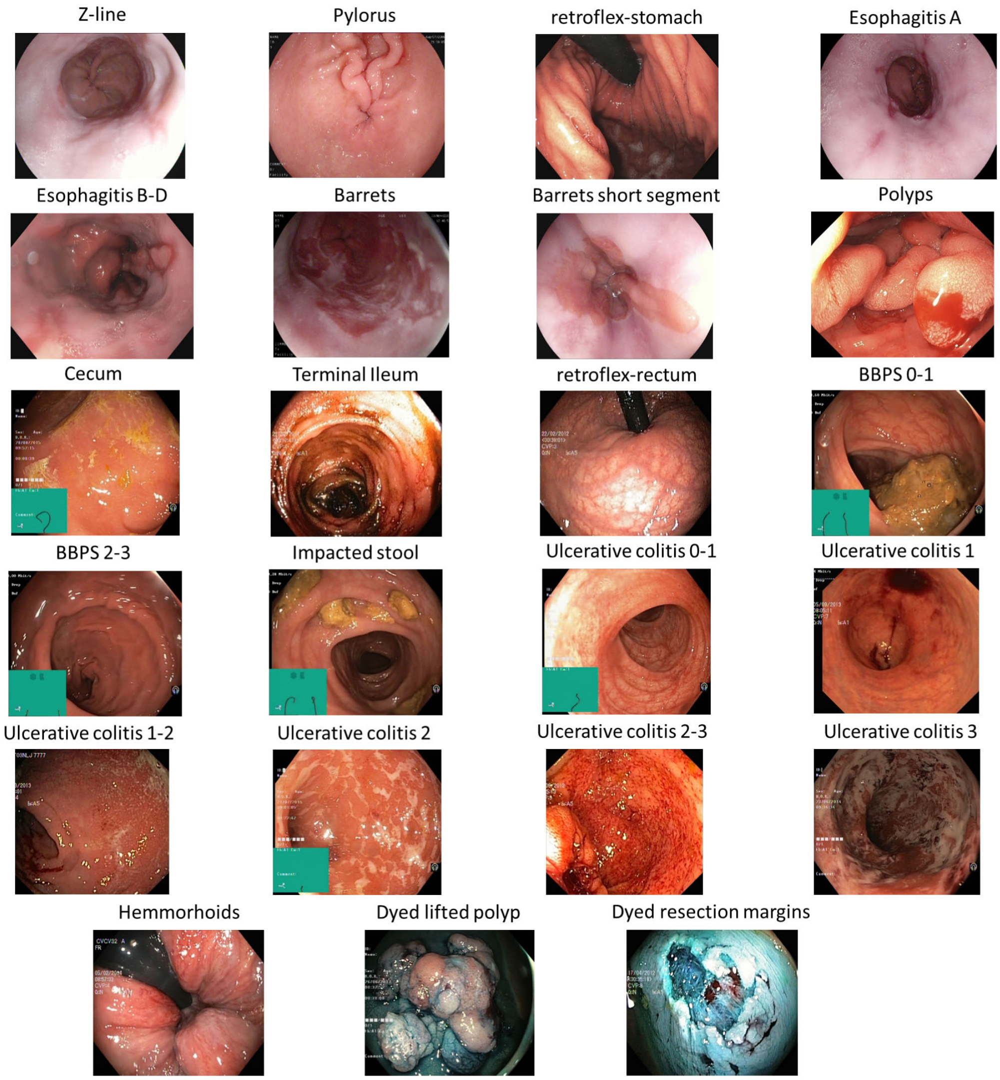

FIGURE 9. Hyper-KVASIR dataset's classes and sample images.

TABLE 8. Details of BioMedia 2019 video dataset.

Disease	Length	Resolution
Esophagitis	00:51	1920 × 1072
Stool	00:02	1920 × 1072
Polyp resection, bleeding	02:00	720 × 576
Bleeding ulcer, instrument	01:08	1280 × 1024
Polyp, lifting and resection, instrument	05:11	720 × 576
Normal colon	00:57	720 × 576

in the medical area due to the fact that certain findings occur more frequently than others. This presents researchers with an additional problem, as AI algorithms applied to the data should also be able to learn from a limited amount of training data.

TABLE 9. Details of hyper-KVASIR dataset.

Data Type	Number of Files	Description
Labeled images	10,662 images	23 different classes
Unlabeled images	99,417 images	unlabeled
Videos	374 videos	30 different classes

F. CLINICAL IMPACT OF AI IN GI ENDOSCOPY AND WIRELESS CAPSULE ENDOSCOPY INVESTIGATION

Given its potential to adapt and continuously learn in real-time, AI-based technology can present unique challenges. Regulatory pathways differ globally, so this review will focus on perspectives from the United States of America (USA), the European Union (EU), and Japan. It also should be noted that a voluntary group known as the International Medical

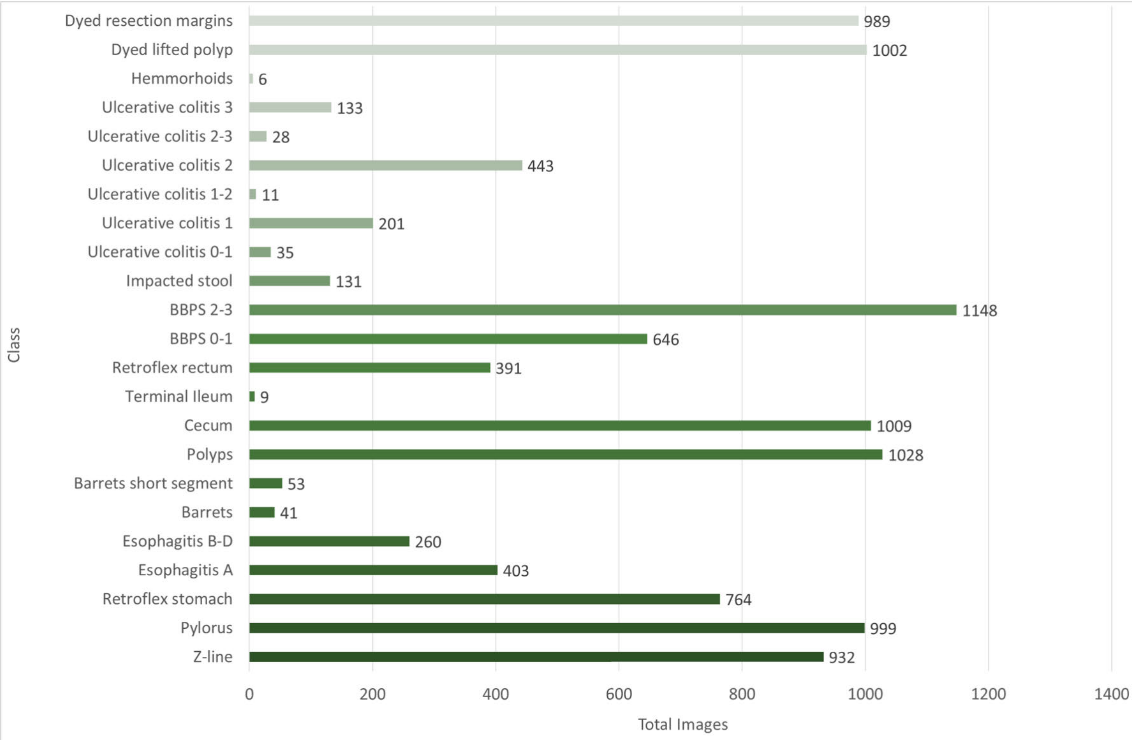

FIGURE 10. Hyper-KVASIR dataset class distribution.

TABLE 10. Details of KVASIR Capsule dataset class labels.

Disease	Class	Description
Ampulla of Vater	Class 0	The junction between the duodenum and the gall duct.
Angiectasia	Class 1	Small superficial dilated vessels causing chronic bleeding and subsequently anaemia.
Blood - fresh	Class 2	Lesions in the upper GI tract or small intestine may bleed, causing the liquid to appear reddened with new blood.
Blood - hematin	Class 3	Minimal bleeding, small black lines observed on the surface of the mucosa.
Erosion	Class 4	The surface of the mucosa is eroded to differing degrees by excavated lesions. Most prevalent are erosions coated by a thin layer of fibrin.
Erythema	Class 5	Changes to the mucosal surface. The mucosal changes emerge as a reddish appearance.
Foreign body	Class 6	Tablet residue or retained capsules, which can also be observed in the lumen.
Ileocecal valve	Class 7	The transition from the small bowels to the large bowel.
Lymphangiectasia	Class 8	Represents dilated lymphoid vessels in the mucosal wall.
Normal clean mucosa	Class 9	Depicts clean small bowel with no or small amount of fluid and mucosa with healthy villi and no pathological finding.
Polyp	Class 10	Can be precancerous lesions, as they are visible as protruding from the mucosal wall.
Pylorus	Class 11	The anatomical junction between the stomach and small bowels and is a sphincter (circular muscle) regulating the emptying of the stomach into the duodenum.
Reduced mucosal view	Class 12	Shows small bowel content reducing the view of the mucosa, like stools or bubbles.
Ulcer	Class 13	Larger erosions.

Device Regulators Forum (IMDRF) is attempting to develop harmonized principles and common regulatory frameworks for software as a medical device (SaMD) [97]. Currently, AI-based GI endoscopy systems are at the jump-off point from proof-of-principle studies to clinical trials. Table 12 shows several commercial AI that have been applied and approved

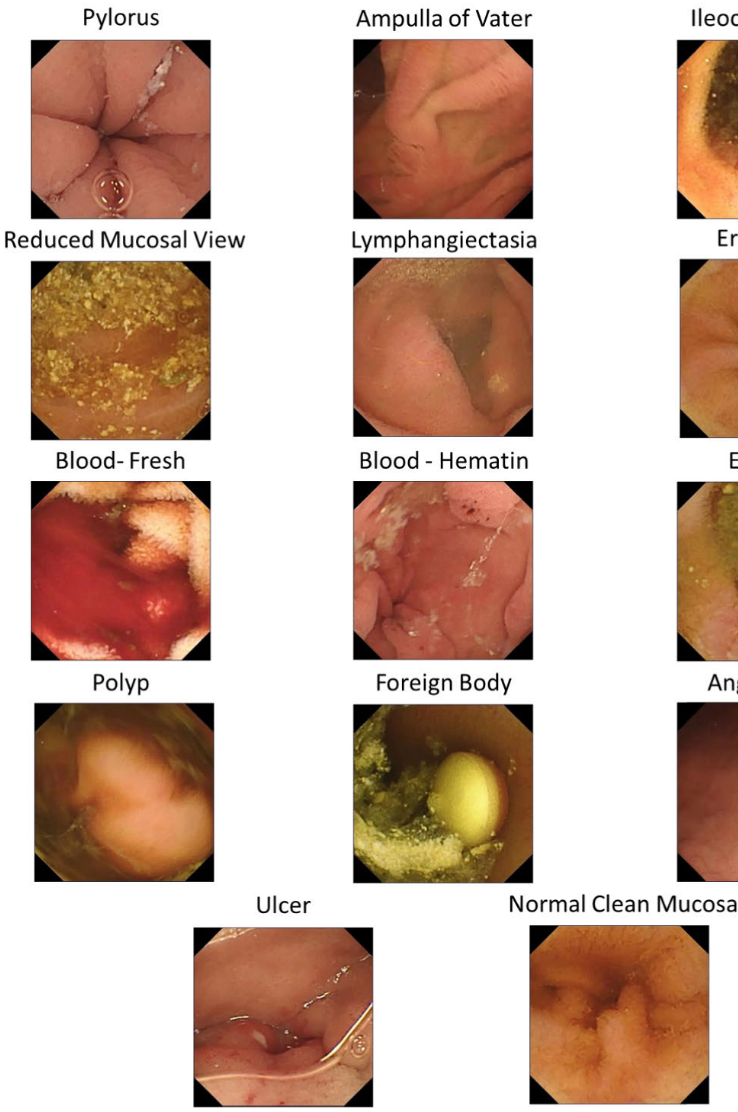

FIGURE 11. KVASIR Capsule dataset's classes and sample images.

TABLE 11. Details of KVASIR Capsule dataset.

Data Type	Number of Files	Description
Labeled images	47,238 images	14 different classes.
Labeled videos	43 videos	Approximately 19 hours of videos; 1,955,675 video frames.
Unlabeled images	1,932,047 images	Unlabeled.
Unlabeled videos	74 videos	Approximatley 25 hours of videos; 2,785,829 video frames.

to use in GI. The EndoBRAIN and EndoBRAIN-UC use a ML algorithm (e.g., SVM) for their systems. The CAD EYE and EndoBRAIN-Eye use deep learning (e.g., CNN) for their systems, while Discovery uses DNN in its system.

Repici et al. [102] utilize GI Genius in a multicenter randomized experiment conducted using 685 subjects in order to compare AI-based GI endoscopy systems to traditional endoscopy. Using GI Genius, they determine that the ADR is significantly higher (i.e., 0.55) compared to traditional endoscopy (i.e., 0.40), and the number of adenomas per colonoscopy is also considerably higher in the AI-based GI endoscopy than in the traditional endoscopy [102]. Additionally, AI-based GI endoscopy systems have faster reaction times compared with endoscopists in 82% of cases with 0.99 sensitivity [99].

Although many of the AI systems listed above sound promising for the future of daily endoscopic applications, the optimum efficacy of AI systems will continue to rely significantly on endoscopists' technical competence. Advanced imaging techniques are also important for capturing stable endoscopic images, which are frequently used as the foundation of AI systems. Wireless capsule endoscopy (WCE) is a

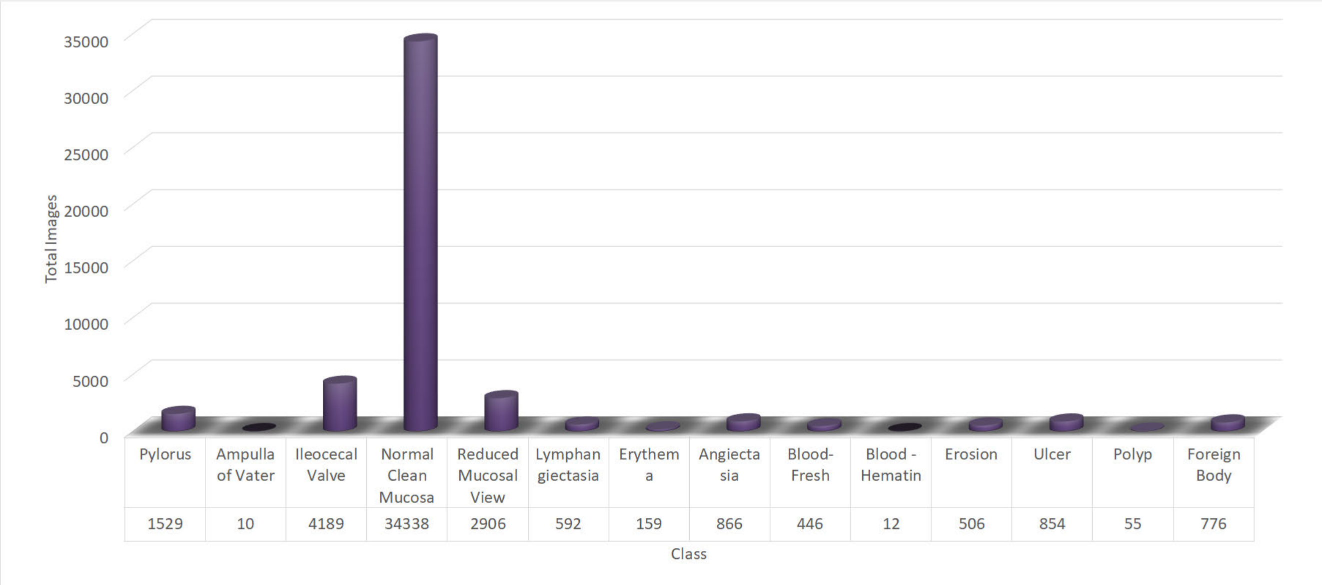

FIGURE 12. KVASIR Capsule dataset class distribution.

TABLE 12. AI applications in GI endoscopy.

Name	Year	State of Approval	Manufacturer	Intended Use	Remarks
AI4GI	2018	CE - EU	Olympus America – Satis Operations and Imagia	Polyp detection and differentiation	Accuracy of 0.97
IRIS	2018	FDA	NinePoint Medical Incorporation	Esophageal image feature segmentation	-
EndoBrain	2018	PMDA	Olympus and Cybernet System Corporation	Neoplastic and non-neoplastic lesions classification	0.97 sensitivity and 0.98 accuracy versus non-specialist clinicians with performance of 0.71 sensitivity and 0.69 accuracy[98]
GI Genius	2019	CE - EU and FDA (2021)	Medtronic	Polyp Detection	14% absolute increase in adenoma detection rate; 0.99 sensitivity rate and less than 0.01 false positives[99]
Discovery	2020	CE - EU	Pentax Medical Corporation	Polyp Detection	120,000 files from approximately 300 clinical cases were used for the software training. No clinical trial report.
Endo-AID	2020	CE - EU	Olympus	Neoplastic and non-neoplastic lesions classification	-
CAD EYE	2020	CE - EU snd FDA	Fujifil GmbH	Polyp Detection	-
EndoBrain-EYE	2020	PMDA	Cybernet System Corporation	Lesion detection	0.90 sensitivity and 0.94 specificity[100]
EndoBrain-UC	2020	PMDA	Olympus and Cybernet System Corporation	Ulcerative colitis (UC) detection	0.74 sensitivity and 0.97 specificity[101]

CE, Conformité Européenne; EU, European Union; FDA, Food and Drug Administration, United States of America; GmbH, Gesellschaft mit beschränkter Haftung; IRIS, Intelligent Real-time Image Segmentation™; PMDA, Pharmaceuticals and Medical Devices Agency, Japan

relatively new technique that allows doctors (i.e., gastroenterologists) to examine the GI interior using a noninvasive procedure. Prior to the introduction of WCE, it had been impossible for a physician to examine GI tissues without performing a surgical procedure. Although WCE has the advantage of investigating the entire digestive system, viewing and evaluating each WCE video takes time and poses numerous challenges. Thus, combining the current AI systems' good diagnostic performances for GI endoscopy with well-trained endoscopists and recent imaging techniques (e.g., WCE) may improve daily endoscopy quality. Researchers in the medical, computer vision, and robotic fields examine GI analysis extensively, and they hope to utilize AI in WCE [103]. Although AI has made significant advancements in GI endoscopy, its contribution to improved WCE is still lacking. Therefore, more research is required to advance AI systems for GI endoscopy to the next development stage for implementation in WCE. Future technologies will face more rigorous approval procedures. Additionally, it is critical to emphasize that medical specialists who provide the necessary expertise in a specific subject play the most essential function in developing an AI medical system.

V. CONCLUSION

This study provides a comprehensive overview of ML and DL algorithms for AI applications in capsule endoscopy, with a particular emphasis on gastrointestinal (GI) classification and detection. In this research, previous studies were analyzed and categorized according to the ML and DL models. In addition, some of the current research issues that are being faced were discussed, which offered researchers new perspectives on the most recent applications of AI in capsule endoscopy. In conclusion, this section reviews the responses to the questions raised in the previous section, Section II.

This study discovered that for the ML model, the SVM model is commonly used in GI image processing and has shown accuracy, sensitivity, and specificity ranging from 0.87 to 0.98, 0.85 to 0.98, and 0.93 to 0.98, respectively. On the other hand, the DL model mainly utilized the CNN-based supervised learning object detection models, such as SSD and Mask RCNN in GI image analysis. A novel encoder-decoder semantic segmentation network called the DSRD-Net has also been proposed for segmenting surgical instruments in minimally invasive surgery, which may be useful for future GI research.

This study also discovered that RGB is still the most commonly used picture modality for GI categorization and detection because the color is one of the most essential factors in detecting bleeding locations. However, there is a need to extend and improve the public dataset database for GI classification because this study found that only five public datasets from 2018-2019 are extensively used for AI applications in GI analysis. However, each dataset has its own classes, some of which are absent from the other dataset. Therefore, it is difficult to combine all datasets to address the problem of insufficient data to train a new data-hungry DL model. There is also a need to create a standardized database to hold datasets for the AI-based GI endoscopy system.

Future and ongoing efforts in software methods (i.e., AI) described in this survey should improve image processing and computer vision techniques, but further improvements may not be possible without hardware advancements and medical doctors' participation. Future capsule generations will provide more quality information to software engineers, allowing them to create smarter software systems that can handle unsolved issues and provide more capabilities.

ACKNOWLEDGMENT

(Malinda Vania and Bayu Adhi Tama contributed equally to this work.)

REFERENCES

[1] A. R. Eliakim, "Video capsule endoscopy of the small bowel (PillCam SB)," *Current Opinion Gastroenterol.*, vol. 22, no. 2, pp. 124–127, Mar. 2006.

[2] L. A. de Souza, C. Palm, R. Mendel, C. Hook, A. Ebigbo, A. Probst, H. Messmann, S. Weber, and J. P. Papa, "A survey on Barrett's esophagus analysis using machine learning," *Comput. Biol. Med.*, vol. 96, pp. 203–213, May 2018.

[3] Y. J. Yang and C. S. Bang, "Application of artificial intelligence in gastroenterology," *World J. Gastroenterol.*, vol. 25, no. 14, pp. 1666–1683, Apr. 2019.

[4] W. Du, N. Rao, D. Liu, H. Jiang, C. Luo, Z. Li, T. Gan, and B. Zeng, "Review on the applications of deep learning in the analysis of gastrointestinal endoscopy images," *IEEE Access*, vol. 7, pp. 142053–142069, 2019.

[5] T. Rahim, M. A. Usman, and S. Y. Shin, "A survey on contemporary computer-aided tumor, polyp, and ulcer detection methods in wireless capsule endoscopy imaging," *Computerized Med. Imag. Graph.*, vol. 85, Oct. 2020, Art. no. 101767.

[6] D. Jha, S. Ali, S. Hicks, V. Thambawita, H. Borgli, P. H. Smedsrud, T. de Lange, K. Pogorelov, X. Wang, P. Harzig, and M. T. Tran, "A comprehensive analysis of classification methods in gastrointestinal endoscopy imaging," *Med. Image Anal.*, vol. 70, May 2021, Art. no. 102007.

[7] S. H. Kim and Y. J. Lim, "Artificial intelligence in capsule endoscopy: A practical guide to its past and future challenges," *Diagnostics*, vol. 11, no. 9, p. 1722, Sep. 2021.

[8] P. Muruganantham and S. M. Balakrishnan, "A survey on deep learning models for wireless capsule endoscopy image analysis," *Int. J. Cognit. Comput. Eng.*, vol. 2, pp. 83–92, Jun. 2021.

[9] R. Trasolini and M. F. Byrne, "Artificial intelligence and deep learning for small bowel capsule endoscopy," *Digestive Endoscopy*, vol. 33, no. 2, pp. 290–297, Jan. 2021.

[10] P. Visaggi, N. de Bortoli, B. Barberio, V. Savarino, R. Oleas, E. M. Rosi, S. Marchi, M. Ribolsi, and E. Savarino, "Artificial intelligence in the diagnosis of upper gastrointestinal diseases," *J. Clin. Gastroenterol.*, vol. 56, no. 1, pp. 23–35, 2022.

[11] International Agency of Research on Cancer, *Cancer Today*, World Health Organization. Geneva, Switzerland, 2020. [Online]. Available: http://gco.iarc.fr/today/home

[12] M. B. Amin, F. L. Greene, S. B. Edge, C. C. Compton, J. E. Gershenwald, R. K. Brookland, L. Meyer, D. M. Gress, D. R. Byrd, and D. P. Winchester, "The eighth edition AJCC cancer staging manual: continuing to build a bridge from a population-based to a more 'personalized' approach to cancer staging," *CA, Cancer J. for Clinicians*, vol. 67, no. 2, pp. 93–99, 2017.

[13] L. Aabakken, A. N. Barkun, P. B. Cotton, E. Fedorov, M. A. Fujino, E. Ivanova, S.-E. Kudo, K. Kuznetzov, T. de Lange, K. Matsuda, O. Moine, B. Rembacken, J.-F. Rey, J. Romagnuolo, T. Rösch, M. Sawhney, K. Yao, and J. D. Waye, "Standardized endoscopic reporting," *J. Gastroenterol. Hepatology*, vol. 29, no. 2, pp. 234–240, Feb. 2014.

[14] B. Doleman, O. Mathiesen, J. C. Jakobsen, A. J. Sutton, S. Freeman, J. N. Lund, and J. P. Williams, "Methodologies for systematic reviews with meta-analysis of randomised clinical trials in pain, anaesthesia, and perioperative medicine," *Brit. J. Anaesthesia*, vol. 126, no. 4, pp. 903–911, Apr. 2021.

[15] M. L. Rethlefsen, S. Kirtley, S. Waffenschmidt, A. P. Ayala, D. Moher, M. J. Page, J. B. Koffel, and P.-S. Group, "PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews," *J. Med. Library Assoc.*, vol. 109, no. 2, p. 39, Jul. 2021.

[16] A. Liberati, D. G. Altman, J. Tetzlaff, C. Mulrow, P. C. Gøtzsche, and J. P. A. Ioannidis, "The PRISMA statement for reporting systematic and meta-analyses of studies that evaluate interventions: Explanation and elaboration," *PLoS Med.*, vol. 6, no. 7, pp. 1–28, 2009.

[17] D. Moher, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Ann. Internal Med.*, vol. 151, no. 4, p. 264, Aug. 2009.

[18] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Inf. Softw. Technol.*, vol. 64, pp. 1–18, Aug. 2015.

[19] Y. Yuan and M. Q.-H. Meng, "Automatic bleeding frame detection in the wireless capsule endoscopy images," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 1310–1315.

[20] Y. Yuan, J. Wang, B. Li, and M. Q.-H. Meng, "Saliency based ulcer detection for wireless capsule endoscopy diagnosis," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 2046–2057, Oct. 2015.

[21] V. S. Charisis and L. J. Hadjileontiadis, "Use of adaptive hybrid filtering process in Crohn's disease lesion detection from real capsule endoscopy videos," *Healthcare Technol. Lett.*, vol. 3, no. 1, pp. 27–33, Mar. 2016.

[22] X. Jia, L. Cai, J. Liu, W. Dai, and M. Q.-H. Meng, "GI bleeding detection in wireless capsule endoscopy images based on pattern recognition and a MapReduce framework," in *Proc. IEEE Int. Conf. Real-Time Comput. Robot. (RCAR)*, Jun. 2016, pp. 266–271.

[23] O. H. Maghsoudi, M. Alizadeh, and M. Mirmomen, "A computer aided method to detect bleeding, tumor, and disease regions in wireless capsule endoscopy," in *Proc. IEEE Signal Process. Med. Biol. Symp. (SPMB)*, Dec. 2016, pp. 1–6.

[24] Y. Yuan, B. Li, and M. Q.-H. Meng, "Bleeding frame and region detection in the wireless capsule endoscopy video," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 2, pp. 624–630, Mar. 2016.

[25] S. Charfi and M. E. Ansari, "Gastrointestinal tract bleeding detection from wireless capsule endoscopy videos," in *Proc. 2nd Int. Conf. Internet Things, Data Cloud Comput.*, Mar. 2017, pp. 1–5.

[26] S. Charfi and M. El Ansari, "Computer-aided diagnosis system for ulcer detection in wireless capsule endoscopy videos," in *Proc. Int. Conf. Adv. Technol. Signal Image Process. (ATSIP)*, May 2017, pp. 1–5.

[27] S. Suman, F. A. B. Hussin, A. S. Malik, K. Pogorelov, M. Riegler, S. H. Ho, I. Hilmi, and K. L. Goh, "Detection and classification of bleeding region in WCE images using color feature," in *Proc. 15th Int. Workshop Content-Based Multimedia Indexing*, Jun. 2017, pp. 1–6.

[28] Y. Yuan, B. Li, and M. Q.-H. Meng, "WCE abnormality detection based on saliency and adaptive locality-constrained linear coding," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 1, pp. 149–159, Jan. 2017.

[29] F. Deeba, M. Islam, F. M. Bui, and K. A. Wahid, "Performance assessment of a bleeding detection algorithm for endoscopic video based on classifier fusion method and exhaustive feature selection," *Biomed. Signal Process. Control*, vol. 40, pp. 415–424, Feb. 2018.

[30] T. Ghosh, S. A. Fattah, and K. A. Wahid, "CHOBS: Color histogram of block statistics for automatic bleeding detection in wireless capsule endoscopy video," *IEEE J. Transl. Eng. Health Med.*, vol. 6, pp. 1–12, 2018.

[31] T. Ghosh, S. A. Fattah, and K. A. Wahid, "Automatic computer aided bleeding detection scheme for wireless capsule endoscopy (WCE) video based on higher and lower order statistical features in a composite color," *J. Med. Biol. Eng.*, vol. 38, no. 3, pp. 482–496, Jun. 2018.

[32] E. Tuba, S. Tomic, M. Beko, D. Zivkovic, and M. Tuba, "Bleeding detection in wireless capsule endoscopy images using texture and color features," in *Proc. 26th Telecommun. Forum (TELFOR)*, Nov. 2018, pp. 1–4.

[33] X. Xing, X. Jia, and M. Q.-H. Meng, "Bleeding detection in wireless capsule endoscopy image video using superpixel-color histogram and a subspace KNN classifier," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 1–4.

[34] A. K. Kundu and S. A. Fattah, "Probability density function based modeling of spatial feature variation in capsule endoscopy data for automatic bleeding detection," *Comput. Biol. Med.*, vol. 115, Dec. 2019, Art. no. 103478.

[35] A. A. Mamun and M. S. Hossain, "Ulcer detection in image converted from video footage of wireless capsule endoscopy," in *Proc. 1st Int. Conf. Adv. Sci., Eng. Robot. Technol. (ICASERT)*, May 2019, pp. 1–4.

[36] A. A. Mamun, M. S. Hossain, M. M. Hossain, and M. G. Hasan, "Discretion way for bleeding detection in wireless capsule endoscopy images," in *Proc. 1st Int. Conf. Adv. Sci., Eng. Robot. Technol. (ICASERT)*, May 2019, pp. 1–6.

[37] N. Obukhova, A. Motyko, B. Timofeev, and A. Pozdeev, "Method of endoscopic images analysis for automatic bleeding detection and segmentation," in *Proc. 24th Conf. Open Innov. Assoc. (FRUCT)*, Apr. 2019, pp. 285–290.

[38] P. Sivakumar and B. M. Kumar, "A novel method to detect bleeding frame and region in wireless capsule endoscopy video," *Cluster Comput.*, vol. 22, no. S5, pp. 12219–12225, Sep. 2019.

[39] H. Ali, M. Sharif, M. Yasmin, and M. H. Rehmani, "Color-based template selection for detection of gastric abnormalities in video endoscopy," *Biomed. Signal Process. Control*, vol. 56, Feb. 2020, Art. no. 101668.

[40] S. Charfi and M. El Ansari, "A locally based feature descriptor for abnormalities detection," *Soft Comput.*, vol. 24, no. 6, pp. 4469–4481, Mar. 2020.

[41] A. K. Kundu, S. A. Fattah, and K. A. Wahid, "Multiple linear discriminant models for extracting salient characteristic patterns in capsule endoscopy images for multi-disease detection," *IEEE J. Transl. Eng. Health Med.*, vol. 8, pp. 1–11, 2020.

[42] A. Rosenfeld et al., "Development and validation of a risk prediction model to diagnose Barrett's oesophagus (MARK-BE): A case-control machine learning approach," *Lancet Digital Health*, vol. 2, no. 1, pp. 37–48, 2020.

[43] Y. Horie, T. Yoshio, K. Aoyama, S. Yoshimizu, Y. Horiuchi, A. Ishiyama, T. Hirasawa, T. Tsuchida, T. Ozawa, S. Ishihara, Y. Kumagai, M. Fujishiro, I. Maetani, J. Fujisaki, and T. Tada, "Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks," *Gastrointestinal Endoscopy*, vol. 89, no. 1, pp. 25–32, Jan. 2019.

[44] Z. Xu, Y. Tao, Z. Wenfang, L. Ne, H. Zhengxing, L. Jiquan, H. Weiling, D. Huilong, and S. Jianmin, "Upper gastrointestinal anatomy detection with multi-task convolutional neural networks," *Healthcare Technol. Lett.*, vol. 6, no. 6, pp. 176–180, Dec. 2019.

[45] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," in *Computer Vision—ECCV*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham, Switzerland: Springer, 2016, pp. 21–37.

[46] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.

[47] M. A. Khan, M. A. Khan, F. Ahmed, M. Mittal, L. M. Goyal, D. J. Hemanth, and S. C. Satapathy, "Gastrointestinal diseases segmentation and classification based on duo-deep architectures," *Pattern Recognit. Lett.*, vol. 131, pp. 193–204, Mar. 2020.

[48] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[49] K. Pogorelov, O. Ostroukhova, A. Petlund, P. Halvorsen, T. de Lange, H. N. Espeland, T. Kupka, C. Griwodz, and M. Riegler, "Deep learning and handcrafted feature based approaches for automatic detection of angiectasia," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Mar. 2018, pp. 365–368.

[50] Y. Zhang, F. Li, F. Yuan, K. Zhang, L. Huo, Z. Dong, Y. Lang, Y. Zhang, M. Wang, Z. Gao, Z. Qin, and L. Shen, "Diagnosing chronic atrophic gastritis by gastroscopy using artificial intelligence," *Digestive Liver Disease*, vol. 52, no. 5, pp. 566–572, May 2020.

[51] S. Shichijo, S. Nomura, K. Aoyama, Y. Nishikawa, M. Miura, T. Shinagawa, H. Takiyama, T. Tanimoto, S. Ishihara, K. Matsuo, and T. Tada, "Application of convolutional neural networks in the diagnosis of helicobacter pylori infection based on endoscopic images," *EBioMedicine*, vol. 25, pp. 106–111, Nov. 2017.

[52] M. A. Khan, A. Majid, N. Hussain, M. Alhaisoni, Y.-D. Zhang, S. Kadry, and Y. Nam, "Multiclass stomach diseases classification using deep learning features optimization," *Comput., Mater. Continua*, vol. 67, no. 3, pp. 3381–3399, 2021.

[53] S. Van Riel, F. Van Der Sommen, S. Zinger, E. J. Schoon, and P. H. N. de With, "Automatic detection of early esophageal cancer with CNNS using transfer learning," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 1383–1387.

[54] X. Jia and M. Q.-H. Meng, "A deep convolutional neural network for bleeding detection in wireless capsule endoscopy images," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 639–642.

[55] S. Seguí, M. Drozdzal, G. Pascual, P. Radeva, C. Malagelada, F. Azpiroz, and J. Vitria, "Generic feature learning for wireless capsule endoscopy analysis," *Comput. Biol. Med.*, vol. 79, pp. 163–172, Dec. 2016.

[56] H. Chen, X. Wu, G. Tao, and Q. Peng, "Automatic content understanding with cascaded spatial–temporal deep framework for capsule endoscopy videos," *Neurocomputing*, vol. 229, pp. 77–87, Mar. 2017.

[57] X. Jia and M. Q.-H. Meng, "Gastrointestinal bleeding detection in wireless capsule endoscopy images using handcrafted and CNN features," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 3154–3157.

[58] X. Jia and M. Q.-H. Meng, "A study on automated segmentation of blood regions in wireless capsule endoscopy images using fully convolutional networks," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 179–182.

[59] A. K. Sekuboyina, S. T. Devarakonda, and C. S. Seelamantula, "A convolutional neural network approach for abnormality detection in wireless capsule endoscopy," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 1057–1060.

[60] H. Takiyama, T. Ozawa, S. Ishihara, M. Fujishiro, S. Shichijo, S. Nomura, M. Miura, and T. Tada, "Automatic anatomical classification of esophagogastroduodenoscopy images using deep convolutional neural networks," *Sci. Rep.*, vol. 8, no. 1, p. 7497, May 2018.

[61] M. Gadermayr, G. Wimmer, H. Kogler, A. Vécsei, D. Merhof, and A. Uhl, "Automated classification of celiac disease during upper endoscopy: Status quo and quo vadis," *Comput. Biol. Med.*, vol. 102, pp. 221–226, Nov. 2018.

[62] T. Ghosh, L. Li, and J. Chakareski, "Effective deep learning for semantic segmentation based bleeding zone detection in capsule endoscopy images," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 3034–3038.

[63] D. K. Iakovidis, S. V. Georgakopoulos, M. Vasilakakis, A. Koulaouzidis, and V. P. Plagianakos, "Detecting and locating gastrointestinal anomalies using deep learning and iterative cluster unification," *IEEE Trans. Med. Imag.*, vol. 37, no. 10, pp. 2196–2210, Oct. 2018.

[64] A. A. Shvets, V. I. Iglovikov, A. Rakhlin, and A. A. Kalinin, "Angiodysplasia detection and localization using deep convolutional neural networks," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2018, pp. 612–617.

[65] H. Alaskar, A. Hussain, N. Al-Aseem, P. Liatsis, and D. Al-Jumeily, "Application of convolutional neural networks for automated ulcer detection in wireless capsule endoscopy images," *Sensors*, vol. 19, no. 6, p. 1265, Mar. 2019.

[66] S.-L. Cai, B. Li, W.-M. Tan, X.-J. Niu, H.-H. Yu, L.-Q. Yao, P.-H. Zhou, B. Yan, and Y.-S. Zhong, "Using a deep learning system in endoscopy for screening of early esophageal squamous cell carcinoma (with video)," *Gastrointestinal Endoscopy*, vol. 90, no. 5, pp. 745–753, Nov. 2019.

[67] T. Cogan, M. Cogan, and L. Tamil, "MAPGI: Accurate identification of anatomical landmarks and diseased tissue in gastrointestinal tract using deep learning," *Comput. Biol. Med.*, vol. 111, Aug. 2019, Art. no. 103351.

[68] M. Hajabdollahi, R. Esfandiarpoor, P. Khadivi, S. M. R. Soroushmehr, N. Karimi, K. Najarian, and S. Samavi, "Segmentation of bleeding regions in wireless capsule endoscopy for detection of informative frames," *Biomed. Signal Process. Control*, vol. 53, Aug. 2019, Art. no. 101565.

[69] C. Lee, J. Min, J. Cha, and S. Lee, "Feature space extrapolation for ulcer classification in wireless capsule endoscopy images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 100–103.

[70] M. Hajabdollahi, R. Esfandiarpoor, K. Najarian, N. Karimi, S. Samavi, and S. M. R. Soroushmehr, "Low complexity CNN structure for automatic bleeding zone detection in wireless capsule endoscopy imaging," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 7227–7230.

[71] M. A. Khan, M. Sharif, T. Akram, M. Yasmin, and R. S. Nayak, "Stomach deformities recognition using rank-based deep features selection," *J. Med. Syst.*, vol. 43, no. 12, p. 329, Dec. 2019.

[72] E. Klang, Y. Barash, R. Y. Margalit, S. Soffer, O. Shimon, A. Albshesh, S. Ben-Horin, M. M. Amitai, R. Eliakim, and U. Kopylov, "Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy," *Gastrointestinal Endoscopy*, vol. 91, no. 3, pp. 606–613, Mar. 2020.

[73] H. Luo, G. Xu, C. Li, L. He, L. Luo, Z. Wang, B. Jing, Y. Deng, Y. Jin, Y. Li, and B. Li, "Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: A multicentre, case-control, diagnostic study," *Lancet Oncol.*, vol. 20, no. 12, pp. 1645–1654, Dec. 2019.

[74] S. Wang, Y. Xing, L. Zhang, H. Gao, and H. Zhang, "A systematic evaluation and optimization of automatic detection of ulcers in wireless capsule endoscopy on a large dataset using deep convolutional neural networks," *Phys. Med. Biol.*, vol. 64, no. 23, Dec. 2019, Art. no. 235014.

[75] S. Wang, Y. Xing, L. Zhang, H. Gao, and H. Zhang, "Deep convolutional neural network for ulcer recognition in wireless capsule endoscopy: Experimental feasibility and optimization," *Comput. Math. Methods Med.*, vol. 2019, pp. 1–14, Sep. 2019.

[76] X. Xing, Y. Yuan, X. Jia, and M. Q.-H. Meng, "A saliency-aware hybrid dense network for bleeding detection in wireless capsule endoscopy images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 104–107.

[77] W. Yang, Y. Cao, Q. Zhao, Y. Ren, and Q. Liao, "Lesion classification of wireless capsule endoscopy images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1238–1242.

[78] S. Özturk and U. Özkaya, "Gastrointestinal tract classification using improved LSTM based CNN," *Multimedia Tools Appl.*, vol. 79, nos. 39–40, pp. 28825–28840, Oct. 2020.

[79] C. S. Bang, H. Lim, H. M. Jeong, and S. H. Hwang, "Use of endoscopic images in the prediction of submucosal invasion of gastric neoplasms: Automated deep learning model development and usability study," *J. Med. Internet Res.*, vol. 23, no. 4, Apr. 2021, Art. no. e25167.

[80] T. Ghosh and J. Chakareski, "Deep transfer learning for automated intestinal bleeding detection in capsule endoscopy imaging," *J. Digit. Imag.*, vol. 34, no. 2, pp. 404–417, Apr. 2021.

[81] M. H. Al-Adhaileh, E. M. Senan, F. W. Alsaade, T. H. H. Aldhyani, N. Alsharif, A. Abdullah Alqarni, M. I. Uddin, M. Y. Alzahrani, E. D. Alzain, and M. E. Jadhav, "Deep learning algorithms for detection and classification of gastrointestinal diseases," *Complexity*, vol. 2021, pp. 1–12, Oct. 2021.

[82] E. Klang, Y. Barash, A. Levartovsky, N. B. Lederer, and A. Lahat, "Differentiation between malignant and benign endoscopic images of gastric ulcers using deep learning," *Clin. Experim. Gastroenterol.*, vol. Volume 14, pp. 155–162, May 2021.

[83] J. Yogapriya, V. Chandran, M. G. Sumithra, P. Anitha, P. Jenopaul, and C. S. G. Dhas, "Gastrointestinal tract disease classification from wireless endoscopy images using pretrained deep learning model," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–12, Sep. 2021.

[84] M. Xu, W. Zhou, L. Wu, J. Zhang, J. Wang, G. Mu, X. Huang, Y. Li, J. Yuan, Z. Zeng, Y. Wang, L. Huang, J. Liu, and H. Yu, "Artificial intelligence in the diagnosis of gastric precancerous conditions by image-enhanced endoscopy: A multicenter, diagnostic study (with video)," *Gastrointestinal Endoscopy*, vol. 94, no. 3, pp. 540–548, Sep. 2021.

[85] V. Siripoppohn, R. Pittayanon, K. Tiankanon, N. Faknak, A. Sanpavat, N. Klaikaew, P. Vateekul, and R. Rerknimitr, "Real-time semantic segmentation of gastric intestinal metaplasia using a deep learning approach," *Clin. Endoscopy*, vol. 55, no. 3, pp. 390–400, May 2022.

[86] M. Navarro, J. Ruberte, A. Carretero, V. Nacher, and E. Domínguez, "5—Digestive tract," in *Morphological Mouse Phenotyping*, J. Ruberte, A. Carretero, and M. Navarro, Eds. Cambridge, MA, USA: Academic Press, 2017, pp. 89–146.

[87] V. Mahadevan, "Anatomy of the stomach," *Surgery*, vol. 32, no. 11, pp. 571–574, 2014.

[88] C. Spada, M.-E. Riccioni, R. Urgesi, and G. Costamagna, "Capsule endoscopy in celiac disease," *World J. Gastroenterol.*, vol. 14, no. 26, pp. 4146–4151, Jul. 2008.

[89] E. Redondo-Cerezo, A. D. Sanchez-Capilla, P. De La Torre-Rubio, and J. De Teresa, "Wireless capsule endoscopy: Perspectives beyond gastrointestinal bleeding," *World J. Gastroenterol.*, vol. 20, no. 42, p. 15664, 2014.

[90] F. Busti, N. Campostrini, N. Martinelli, and D. Girelli, "Iron deficiency in the elderly population, revisited in the hepcidin era," *Frontiers Pharmacol.*, vol. 5, p. 83, Apr. 2014.

[91] Q. Zhao and T. Chi, "Deep learning model can improve the diagnosis rate of endoscopic chronic atrophic gastritis: A prospective cohort study," *BMC Gastroenterol.*, vol. 22, no. 1, p. 133, Mar. 2022, doi: 10.1186/s12876-022-02212-1.

[92] K. Pogorelov, K. R. Randel, C. Griwodz, S. L. Eskeland, T. de Lange, D. Johansen, C. Spampinato, D.-T. Dang-Nguyen, M. Lux, P. T. Schmidt, M. Riegler, and P. Halvorsen, "KVASIR: A multi-class image dataset for computer aided gastrointestinal disease detection," in *Proc. 8th ACM Multimedia Syst. Conf.* New York, NY, USA: Association for Computing Machinery, 2017, pp. 164–169.

[93] K. Pogorelov, M. Riegler, P. Halvorsen, S. Hicks, K. R. Randel, D. T. D. Nguyen, M. Lux, O. Ostroukhova, and T. de Lange, "Medico multimedia task at mediaeval 2018," in *Proc. CEUR Workshop*, vol. 2283. Aachen, Germany: Technical Univ. Aachen, 2018, pp. 1–4.

[94] S. Hicks, M. Riegler, P. Smedsrud, T. B. Haugen, K. R. Randel, K. Pogorelov, H. K. Stensland, D.-T. Dang-Nguyen, M. Lux, A. Petlund, T. de Lange, P. T. Schmidt, and P. Halvorsen, "ACM multimedia BioMedia 2019 grand challenge overview," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 2563–2567.

[95] H. Borgli, V. Thambawita, P. H. Smedsrud, S. Hicks, D. Jha, S. L. Eskeland, K. R. Randel, K. Pogorelov, M. Lux, D. T. D. Nguyen, D. Johansen, C. Griwodz, H. K. Stensland, E. Garcia-Ceja, P. T. Schmidt, H. L. Hammer, M. A. Riegler, P. Halvorsen, and T. de Lange, "HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy," *Sci. Data*, vol. 7, no. 1, p. 283, Aug. 2020.

[96] P. H. Smedsrud et al., "*Kvasir-Capsule*, a video capsule endoscopy dataset," *Sci. Data*, vol. 8, no. 1, p. 142, 2021.

[97] N. Aisu, M. Miyake, K. Takeshita, M. Akiyama, R. Kawasaki, K. Kashiwagi, T. Sakamoto, T. Oshika, and A. Tsujikawa, "Regulatory-approved deep learning/machine learning-based medical devices in Japan as of 2020: A systematic review," *PLOS Digit. Health*, vol. 1, no. 1, pp. 1–12, 2022.

[98] S. E. Kudo et al., "Artificial intelligence-assisted system improves endoscopic identification of colorectal neoplasms," *Clin. Gastroenterol. Hepatology*, vol. 18, no. 8, pp. 1874–1881, Jul. 2020.

[99] C. Hassan, M. B. Wallace, P. Sharma, R. Maselli, V. Craviotto, M. Spadaccini, and A. Repici, "New artificial intelligence system: First validation study versus experienced endoscopists for colorectal polyp detection," *Gut*, vol. 69, no. 5, pp. 799–800, May 2020.

[100] M. Misawa, S.-E. Kudo, Y. Mori, K. Hotta, K. Ohtsuka, T. Matsuda, S. Saito, T. Kudo, T. Baba, F. Ishida, H. Itoh, M. Oda, and K. Mori, "Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video)," *Gastrointestinal Endoscopy*, vol. 93, no. 4, pp. 960–967, Apr. 2021.

[101] Y. Maeda, S.-E. Kudo, Y. Mori, M. Misawa, N. Ogata, S. Sasanuma, K. Wakamura, M. Oda, K. Mori, and K. Ohtsuka, "Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video)," *Gastrointestinal Endoscopy*, vol. 89, no. 2, pp. 408–415, Feb. 2019.

[102] A. Repici et al., "Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial," *Gastroenterology*, vol. 159, no. 2, pp. 512–520, Aug. 2020.

[103] J. Song, M. Patel, A. Girgensohn, and C. Kim, "Combining deep learning with geometric features for image-based localization in the gastrointestinal tract," *Exp. Syst. Appl.*, vol. 185, Dec. 2021, Art. no. 115631. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417421010253

MALINDA VANIA received the Ph.D. degree from the Korea University of Science and Technology (UST) and the Korea Institute of Science and Technology (KIST), Republic of Korea. She was an Assistant Researcher with the Center for Healthcare Robotics, KIST, Seoul, Republic of Korea. She is currently a Postdoctoral Researcher with the Unstructured Data Mining and Machine Learning Laboratory, Department of Industrial Engineering, Ulsan National Institute of Science and Technology (UNIST), Republic of Korea. Her research interests include applied deep learning for healthcare, medical image processing, image registration, image segmentation, image synthesis, explainable artificial intelligence, and sound analysis.

BAYU ADHI TAMA (Member, IEEE) was a Postdoctoral Researcher with the Pohang University of Science and Technology and the Ulsan National Institute of Science and Technology, Republic of Korea, in 2020 and 2018, respectively. He was also a Senior Research Fellow with the Data Science Group, Institute for Basic Science (IBS), in 2022. He is currently a Research Scientist with the NSF HDR Institute for Harnessing Data and Model Revolution in the Polar Regions (iHARP), University of Maryland, Baltimore County (UMBC), Baltimore, MD, USA. His research interests include data-driven computational approaches for healthcare, manufacturing, cybersecurity, computational social science, and data-driven discoveries in polar regions.

HASAN MAULAHELA received the M.D. degree in gastroenterohepatology from the University of Indonesia, Jakarta. He is currently the Head of the Gastrointestinal Endoscopy Unit, Universitas Indonesia-Cipto Mangunkusumo National General Hospital. He is also an Academic Staff with the Division of Gastroenterology, Pancreatobiliary, and Digestive Endoscopy, Department of Internal Medicine, University of Indonesia. He previously trained in advanced endoscopy with the Asian Institute of Gastroenterology, Hyderabad, India, Kobe University, Japan, and the Asan Hospital, Seoul, South Korea. His research interests include Helicobacter pylori and intestinal tuberculosis infection, advanced endoscopy techniques, simulation-based learning, and clinical application of artificial intelligence (AI) in GI endoscopy.

SUNGHOON LIM received the B.S. and M.S. degrees in industrial engineering from KAIST, Republic of Korea, in 2005 and 2009, respectively, and the Ph.D. degree in industrial engineering from the Pennsylvania State University, University Park, PA, USA, in 2018. He is currently an Associate Professor with the Department of Industrial Engineering and the Graduate School of Artificial Intelligence and the Head of the Industry Intelligentization Institute, Ulsan National Institute of Science and Technology (UNIST), Republic of Korea. His research interests include machine learning/deep learning, industrial artificial intelligence, and smart manufacturing.