Comprehensive review of publicly available colonoscopic imaging databases for artificial intelligence research: availability, accessibility, and usability

Britt B. S. L. Houwen, MD,[1] Karlijn J. Nass, MD,[1] Jasper L. A. Vleugels, MD, PhD,[1] Paul Fockens, MD, PhD,[1] Yark Hazewinkel, MD, PhD,[2] Evelien Dekker, MD, PhD[1]

Amsterdam, Nijmegen, the Netherlands

Background and Aims: Publicly available databases containing colonoscopic imaging data are valuable resources for artificial intelligence (AI) research. Currently, little is known regarding the available number and content of these databases. This review aimed to describe the availability, accessibility, and usability of publicly available colonoscopic imaging databases, focusing on polyp detection, polyp characterization, and quality of colonoscopy.

Methods: A systematic literature search was performed in MEDLINE and Embase to identify AI studies describing publicly available colonoscopic imaging databases published after 2010. Second, a targeted search using Google's Dataset Search, Google Search, GitHub, and Figshare was done to identify databases directly. Databases were included if they contained data about polyp detection, polyp characterization, or quality of colonoscopy. To assess accessibility of databases, the following categories were defined: open access, open access with barriers, and regulated access. To assess the potential usability of the included databases, essential details of each database were extracted using a checklist derived from the Checklist for Artificial Intelligence in Medical Imaging.

Results: We identified 22 databases with open access, 3 databases with open access with barriers, and 15 databases with regulated access. The 22 open access databases contained 19,463 images and 952 videos. Nineteen of these databases focused on polyp detection, localization, and/or segmentation; 6 on polyp characterization, and 3 on quality of colonoscopy. Only half of these databases have been used by other researcher to develop, train, or benchmark their AI system. Although technical details were in general well reported, important details such as polyp and patient demographics and the annotation process were under-reported in almost all databases.

Conclusions: This review provides greater insight on public availability of colonoscopic imaging databases for AI research. Incomplete reporting of important details limits the ability of researchers to assess the usability of current databases. (Gastrointest Endosc 2023;97:184-99.)

(footnotes appear on last page of article)

The successful development of artificial intelligence (AI) systems for automatic polyp detection, polyp characterization, and colonoscopy quality assessment requires availability of high-quality databases of colonoscopy images or videos along with high-quality annotations.[1-4] This may contribute to a reduced number of missed polyps and improved accuracy of optical diagnosis of colorectal polyps, leading to optimized treatment decisions and potential implementation of the optical diagnosis strategy.[5-7] These annotations provide the ground truth (eg, accurate polyp localization, polyp histology class, or bowel cleanliness level) and are necessary to train machine learning models. For object detection and classification in general, large databases are publicly available and widely used to develop and improve machine learning models.[8] In health care, however, it is difficult to create such large-scale and high-quality publicly available databases because of high monetary costs, required expertise, and privacy regulations. These databases certainly exist, as most gastroenterology departments own clinical colonoscopy imaging data. Because of barriers in accessibility and usability, however, these databases might not be used often by researchers.

Various barriers of accessibility exist: visibility barriers (eg, no centralized directory of colonoscopic imaging databases for AI-related research), cost barriers (eg, payment for access to databases), time barriers, privacy barriers

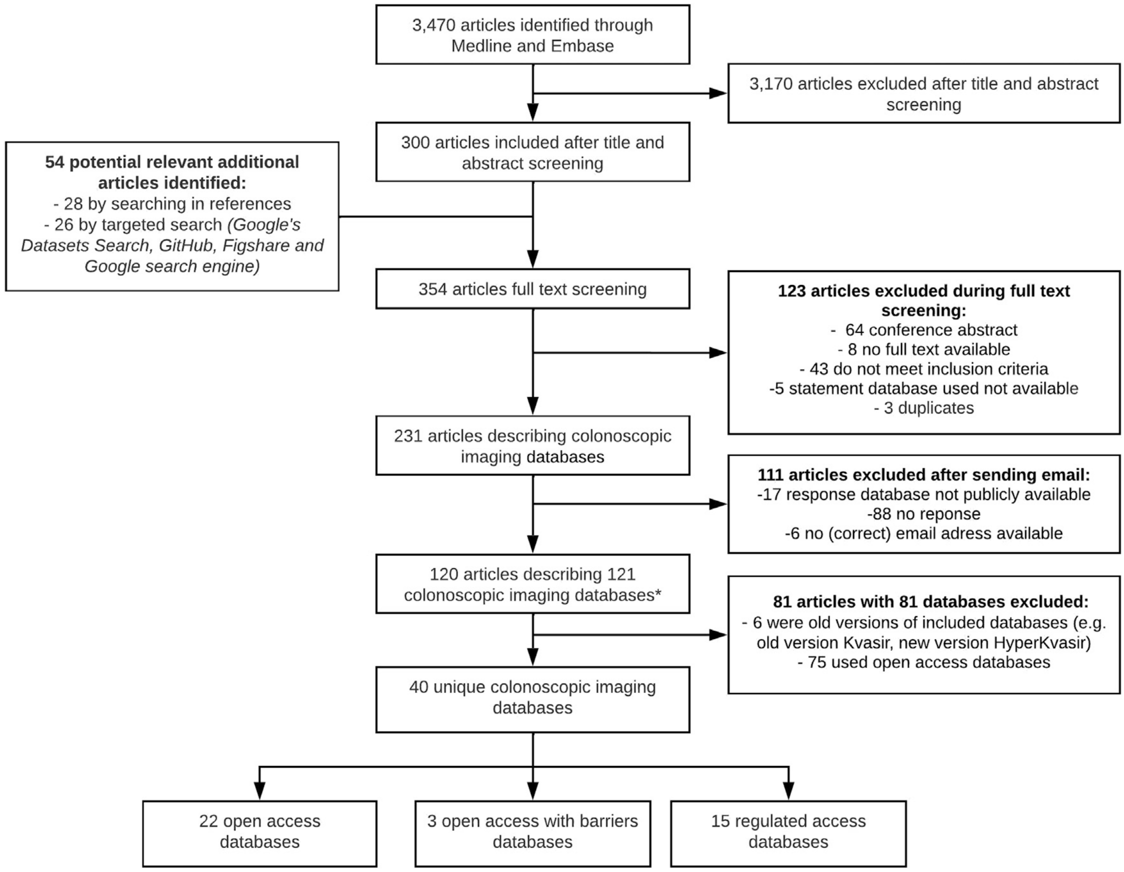


Figure 1. Flowchart of literature and targeted search and database identification. *Wang et al[14] contained 2 databases.

(eg, participant consent, ethical approval), political barriers (eg, researchers or commercial parties do not want to share their data), and governance barriers (eg, regulations on data use). Furthermore, several barriers of usability are imaginable: data quality barriers (eg, low resolution, blurry, "far away" images), data modality barriers (eg, an outdated image modality is used), data representativeness barriers (eg, unrepresentative, unbalanced, or beyond clinical standard high-quality database), and annotation process barriers (eg, no precise documentation of the annotation process). To maximize the potential of data resources and ultimately improve the quality of AI research, this review aimed to identify the availability, accessibility, and usability of publicly available colonoscopic imaging databases focusing on polyp detection, polyp characterization, and quality of colonoscopy.

METHODS

Eligibility criteria and search

This review was performed in compliance with the PRISMA 2020 guidelines[9] and was registered with PROSPERO (CRD42021230672). We included AI-related studies describing colonoscopy imaging databases containing both images and videos. A colonoscopic imaging database was eligible for inclusion if the database contained data about polyp detection (including polyp localization and segmentation), polyp characterization, and/or colonoscopy quality (eg, quality of bowel preparation, cecal intubation landmarks, rectum in retroflexion) (Appendix 1, available online at www.giejournal.org). Only studies in English published since 2010 were included. Studies with databases containing noncolonoscopy images or videos, text, or numeric-only data and images from nonhumans were excluded. Studies focusing on CT colonography, wireless capsule endoscopy, endocytomicroscopy, confocal laser endoscopy, confocal microscopy, or high-resolution microendoscopy were excluded for this review. Studies exclusively focusing on patients with inflammatory bowel disease or polyposis syndromes were excluded as well.

Search strategy

The search consisted of 2 strategies. First, we performed a systematic literature search in MEDLINE and Embase to identify AI studies describing colonoscopic imaging databases that were publicly available, and then we attempted to access these databases at the source (Appendix 2, available online at www.giejournal.org). Second, we performed a targeted

TABLE 1. Characteristics of open access databases and their applications

Database	Country of origin	Publication date	Access type	Data access details	Link to database
BKAI-IGH NeoPolyp-Small[12]	Vietnam	2021	OA	Downloadable zip file	https://bkai.ai/research/bkai-igh-neopolyp-small-a-database-for-fine-grained-polyp-segmentation/
Cho et al 2019[13,‡]	Korea	2019	OA	Downloadable zip file	https://doi.org/10.6084/m9.figshare.7937336.v1
CP-CHILD-A[14]	China	2020	OA	Downloadable zip file	https://doi.org/10.6084/m9.figshare.12554042
CP-CHILD-B[14]	China	2020	OA	Downloadable zip file	https://doi.org/10.6084/m9.figshare.12554042
CVC-ClinicVideoDB[15,16]	Spain	2017	OA	Register and create account for link to download rar file	https://giana.grand-challenge.org/
CVC-EndoSceneStill[17,§]	Spain	2017	OA	Register and create account for link to download zip file	http://www.cvc.uab.es/CVC-Colon/index.php/databases/cvc-endoscenestill/
CVC-HDClassif database[18,19]	Spain	2021	OA	Register and create account for link to download zip files	https://giana.grand-challenge.org/
CVC-PolypHD[17]	Spain	2017	OA	Register and create account for link to download zip file	https://giana.grand-challenge.org/
EDD 2020[20,‖]	France, Italy	2021	OA	Downloadable zip file	https://ieee-dataport.org/competitions/endoscopy-disease-detection-and-segmentation-edd2020
ERS[35,‖]	Poland	2022	OAr	Fill in a form and get e-mailed a link to download zip file	https://cvlab.eti.pg.gda.pl/publications/endoscopy-database
Hamlyn Centre Endoscopic Video Database[21]	UK	2016	OA	Fill in a form for link to download zip file	http://hamlyn.doc.ic.ac.uk/vision
HyperKvasir[22,¶]	Norway	2019	OA	Downloadable zip file	https://osf.io/mh9sj/
ISIT-UMR Colonoscopy database[23]	France	2016	OA	Download each file separately	http://www.depeca.uah.es/colonoscopy_database/
KUMC database[24-26]	USA	2020	OA	Downloadable zip file	https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0236452#sec019
LDPolypVideo Benchmark[27]	China	2021	OA	Downloadable zip file	https://github.com/dashishi/LDPolypVideo-Benchmark
NBIPolyp-Ucdb[28]	Portugal	2019	OA	Fill in a form for link to download zip file	http://www.mat.uc.pt/~isabelf/Polyp-UCdb/NBIPolyp-UCdb.html
WLPolyp-UCdb[29]	Portugal	2019	OA	Fill in a form for link to download zip file	http://www.mat.uc.pt/~isabelf/Polyp-UCdb/WLPolyp-UCdb.html
Nerthus[30]	Norway	2017	OA	Downloadable zip file	https://databases.simula.no/nerthus/
PICCOLO Widefield database[31]	Spain	2020	OAr	Fill in a form and get e-mailed a link to download zip file	https://www.biobancovasco.org/en/Sample-and-data-catalog/Databases/PD178-PICCOLO-EN.html
POLAR database[32]	The Netherlands	2022	OAr	Fill in a form and get e-mailed a link to download zip file	http://www.polar.amsterdamumc.org
PolypGen[33]	UK, Norway, Egypt, Italy, France, Sweden	2021	OA	Register and create account for link to download zip file	https://www.synapse.org/; https://www.synapse.org/#!Synapse:syn26376620.1
SUN Colonoscopy video database[34]	Japan	2020	OAr	Request by e-mail for link to download zip file	http://amed8k.sundatabase.org/

OAr, Access on request; *BP*, bowel preparation; *C*, commercial; NR, not reported; NC, noncommercial; OA, open access; PC, polyp characterization; PD, polyp detection; PL, polyp localization; PS, polyp segmentation.

*Prospective database acquisition means the database was prospectively collected with the purpose of training/external validation database.

†Supplementary Table 4 shows the citations that used this database for training and/or external validation.

‡Complete database contains 112 patients; however, data only available for 2 patients.

§Derived from CVC-ColonDB[55] and CVC-ClinicDB.[56]

‖For review only colorectal cancer and polyp data were included.

¶Derived from Kvasir[57] and Kvasir-SEG[58] for review only lower GI data focused on colorectal polyp detection, localization, segmentation, characterization, and quality of colonoscopy were included.

**Not explicitly described whether the database can be used for commercial purposes.

TABLE 1. Continued

Database application	Acquisition data*	No. of centers collecting data	No. of patients	Total no. of images	Total no. of videos(frames)	Purpose reuse	Database reuse (impact)		
							Cited	Training/ internal validation†	Testing/ external validation†
PD, PL, PS, PC	NR	1	NR	1200	0	NR	3	0	0
Cecal landmarks	NR	1	2	0	2 (3645)	NC/C	4	0	0
PD	Prospective	1	NR	8000	0	NC/C	20	1	0
PD	Prospective	1	NR	1500	0	NC/C	20	0	1
PD, PL, PS	NR	1	36	0	36 (29,657)	NC	27	16	4
PD, PL, PS	Prospective	1	36	0	44 (912)	NC	772	165	19
PD, PL, PS, PC	NR	1	NR	1001	0	NC	0	0	0
PD, PL, PS	NR	1	NR	164	0	NC	131	3	0
PD, PL, PS	Prospective	2	NR	0	NR (175)	NR	30	3	0
PD, PL, PS	Retrospective	1	708	0	NR (44,249)	NC	0	0	0
PD, PL	Prospective	NR	NR	0	10 (7894)	NR	43	4	1
PD, PL, PS, BP, cecal landmarks, RF	Prospective	1	NR	4,231	142 (NR)	NC/C**	474	107	4
PC	Prospective	NR	NR	0	152 (NR)	NC/C**	107	28	1
PD, PL, PC	NR	1	NR	0	76 (4955)	NC	30	2	0
PD, PL	NR	1	NR	0	263 (901,666)	NC/C	0	0	0
PD, PL, PS	Prospective	1	10	0	11 (86)	NC	2	0	0
PD	Prospective	1	42	0	42 (3040)	NC	23	0	0
BP	Prospective	1	21	0	21 (5,525)	NC/C**	53	1	1
PD, PL, PS, PC	Prospective	1	40	0	40 (3433)	NC	11	2	0
PD, PL, PC	Prospective	9	802	3367	0	NC	0	0	0
PD, PL, PS	Prospective	6	>300	0	NR (6282)	NC	7	6	0
PD, PL, PC	Prospective	5	99	0	113 (158,690)	NC	48	5	1

search with similar search terms using Google's Database Search, the Google search engine, GitHub, and Figshare to identify colonoscopic imaging databases directly. For both Google search and Google Dataset search, results returned for each search were systematically collated and screened. This systematic search was conducted under supervision of a medical librarian. Records were managed through Rayyan (https://www.rayyan.ai/) and Endnote (https://endnote.com/), which are specific software for managing bibliographies

TABLE 2. Characteristics of open access databases focusing on polyp detection, localization, and segmentation

Database	Application database	No. of unique patients	No. of unique polyps	No. of images				No. of videos (frames)			
				Total	With polyp	Without polyp	Not annotated	Total	With polyp	Without polyp	Not annotated
BKAI-IGH NeoPolyp-Small[12]	PD, PL, PS	NR	6	1200	1,00	0	0	0	0	0	0
CP-CHILD-A[14]	PD	NR	NR	8000*	1000*	7000*	0	0	0	0	0
CP-CHILD-B[14]	PD	NR	NR	1500†	400†	1,100†	0	0	0	0	0
CVC-ClinicVideoDB[15,16]	PD, PL, PS	36	36		0	0	0	36 (29,657)‡	18 (9545)	18 (1384)	18 (18,733)
CVC-EndoSceneStill[17]	PD, PL, PS	36	NR	0	0	0	0	44 (912)	44 (912)	0	0
CVC-HDClassif database[18,9]	PD, PL, PS	NR	NR	1001§	901	0	100	0	0	0	0
CVC-PolypHD[17]	PD, PL, PS	NR	164‖	164‖	56	0	108	0	0	0	0
EDD 2020[20]	PD, PL, PS	NR	NR	0	0	0	0	NR (175)	NR (175)	0	0
ERS[35]	PD, PL, PS	708	NR	0	0	0	0	NR (44,249)	NR (32,218)	NR (12,031)	0
Hamlyn Centre Endoscopic Video Database[21]	PD, PL	NR	10	0	0	0	0	10 (7894)	10 (4219)	10 (3675)	0
HyperKvasir[22]	PD, PL, PS	NR	NR	1028	1028 (1000)**	0	0	73 (NR)	73 (NR)	0	0
KUMC database[24-26]	PD, PL	NR	76	0	0	0	0	76 (4955)	76 (4955)	0	0
LDPolypVideo Benchmark[27]	PD, PL	NR	200	0	0	0	0	263 (901,666)	NR (33,884)	NR (6382)	103 (861,400)
NBIPolyp-Ucdb[28]	PD, PL, PS	10	11	0	0	0	0	11 (86)	11 (86)	0	0
WLPolyp-UCdb[29]	PD	42	42	0	0	0	0	42 (3040)	42 (1680)	NR (1360)	0
PICCOLO Widefield database[31]	PD, PL, PS	40	40	0	0	0	0	40 (3433)	40 (3433)	0	0
POLAR database[32]	PD, PL	802	2069	3367††	3367††	0	0	0	0	0	0
PolypGen[33]	PD, PL, PS	>300	NR	0	0	0	0	NR (6282)	NR (3762)	NR (2520)	0
SUN Colonoscopy video database[34]	PD, PL	99	100	0	0	0	0	113 (158,690)	100 (49,136)	13 (109,554)	0

Combination N-E, Combination no physician and expert physician; *Combination J-E*, combination junior physician and expert physician; *PD*, polyp detection; *PL*, polyp localization; *PS*, polyp segmentation; *NR*, not reported; *WL*, white-light endoscopy; *NBI*, narrow-band imaging; FICE, Fuji Intelligent Color Enhancement; CE, dye-based chromoendoscopy.

*7000 training images (800 polyp, 7200 nonpolyp), 1000 testing images (200 polyp, 800 nonpolyp).

†1100 training images (300 polyp, 800 nonpolyp), 400 testing images (100 polyp, 300 nonpolyp).

‡Training (10,924 frames) test (18,733 frames).

§788 training images, 113 internal validation images, 100 testing images.

‖56 training images, 108 test images.

**1028 images annotated with bounding box; 1000 images annotated with binary mask.

††2637 training images (1339 unique polyps), 730 test images (730 unique polyps).

Selection process

Search results were independently screened in duplicate by 2 reviewers (B.B.S.L.H. and K.J.N.) to identify the name and source of any relevant study with an eligible database. Duplicates were removed, and obvious irrelevant studies were excluded based on title and abstract by the reviewers. Subsequently, full-text screening was performed. Studies were classified in the categories “exclude” if they clearly did not meet the eligibility criteria and “include” for studies with a database that met the eligibility criteria. Where status of availability was unclear and to check whether a database was publicly available, and if so under which conditions, an e-mail was sent to the contact person of the study or database. Per protocol, we allowed a 2-week period for an e-mail response. The Google search engine, a Google Dataset Search, GitHub, and Figshare results were screened by the primary reviewers to directly identify relevant databases. A third reviewer (Y.H.) resolved discrepancies from the initial and full-text screenings.

TABLE 2. Continued

Endoscopy system brand	Imaging modality	Image, frame resolution (pixels)	File format	Source of ground truth annotations	
				Ground truth	Annotation performed by
Fujifilm	WL, FICE	1280 × 959	JPEG	Binary mask	Combination N-E
Olympus	NR	256 × 256	JPEG	Annotated file	Expert physician(s)
Fujifilm	NR	256 × 256	JPEG	Annotated file	Expert physician(s)
Olympus	WL	768 × 576	PNG	Binary mask	NR
NR	WL	500 × 574, 384 × 288	BMP	Binary mask	NR
NR	WL	NR	TIF	Binary mask	NR
NR	WL	1920 × 1080	TIF, BMP	Binary mask	NR
Olympus	WL, CE	NR	JPEG	Bounding box, binary mask	Combination N-E
NR	NR	NR	PNG	Annotated file, polygon-shaped mask	Expert physician(s)
Olympus, Pentax	NBI, WL	640 × 480	PNG	Bounding box	Expert physician(s)
Olympus, Pentax	WL	332 × 352 to 1921 × 1073	JPEG	Bounding box, binary mask	Combination J-E
NR	WL	224 × 224	JPG	Bounding box	Physician(s), experience not specified
NR	WL	560 × 480	JPG	Bounding box	Annotation tool
Olympus	NBI	720 × 576	JPEG	Binary mask	Combination J-E
Olympus	WL	720 × 576	JPEG	Annotated file	NR
Olympus	NBI, WL	854 × 480, 1920 × 1080	PNG	Binary mask	Combination N-E
Olympus	NBI	1126 × 900, 1920 × 1080	PNG	Bounding box	Combination N-E
NR	WL	NR	JPG	Bounding box, binary mask	Expert physician(s)
Olympus	WL	416 × 416	JPEG	Bounding box	Combination N-E

Database accessibility

The definition of data accessibility previously proposed by Khan et al[10] was used for this review. Accessibility of databases was defined as open access, for which there were no requirements or minimum requirements for access (eg, creation of a free account); open access with barriers, which were databases fulfilling the theoretical criteria for open access but were inaccessible because of unpredictable reasons (eg, broken hyperlink); or regulated access, which required the fulfillment of formal agreements, approvals, or payment. For the open access category where the database access required an e-mail request, we again allowed a 2-week period for e-mail response. To assess to which extent the identified open access databases were already accessible to other AI researchers, we searched the citations of the identified open access databases using the Google search engine ("database used for developing/training model," "database used for external validation model").

Database usability

No standardized checklist is currently available to assess the usability of imaging databases for AI research. Whether a database is useful for AI system developers is subjective but is closely associated with the quality and representativeness of that database. This can only be properly assessed if essential details of the data (ie, metadata) have been reported. Therefore, to map the potential usability of the included databases as best as possible, metadata of each database were extracted using a structured checklist for metadata reporting (Supplementary Table 1, available online at www.giejournal.org). This checklist was derived from the Checklist for Artificial Intelligence in Medical Imaging.[11]

TABLE 3. Characteristics of open access databases focusing on polyp characterization

Database	No. of unique patients	No. of unique polyps	No. of images				No. of videos (frames)			
			Total	With polyp	Without polyp	Not annotated	Total	With polyp	Without polyp	Not annotated
CVC-HDClassif database[18,19]	NR	NR	1001*	901	0	100	0	0	0	0
ISIT-UMR Colonoscopy database[23]	NR	76	0	0	0	0	152 (NR)	152 (NR)	0	0
KUMC database[24-26]	NR	76	0	0	0	0	76 (4955)	76 (4955)	0	0
PICCOLO Widefield database[31]	48	71	0	0	0	0	39 (3433)	39 (3433)	0	0
POLAR database[32]	802	2069	3367§	3367§	0	0	0	0	0	0
SUN Colonoscopy video database[34]	99	100	0	0	0	0	113 (158,690)	100 (49,136)	13 (109,554)	0

HP, Hyperplastic polyp; HGD, high-grade dysplasia; LGA, low-grade adenoma; LGD: low-grade dysplasia; HGA, high-grade adenoma; TSA, traditional serrated adenoma; NR, not reported; WL, white-light endoscopy; NBI, Narrow-band imaging.
*788 training, 113 validation, 100 test images.
†Paris endoscopic classification.[52]
‡NBI international colorectal endoscopic (NICE) classification.[7]
§2637 training images (1339 unique polyps), 730 test images (730 unique polyps).

TABLE 4. Characteristics of open databases focusing on quality of colonoscopy

Database	No. of unique patients	No. of unique polyps	Database of images and/or videos	No. of images				No. of videos (frames)			
				Total	Quality of bowel preparation	Cecal intubation landmarks	Rectal retroflexion	Total	Quality of bowel preparation	Cecal intubation landmarks	Rectal retroflexion
Cho et al 2019[13]	2*	NR	Videos	0	0	0	0	2 (3645)	0	2 (3645)	0
HyperKvasir[22]	NR	NR	Combination	4231	1794	1018	391	142 (NR)	65 (NR)	4 (NR)	0
Nerthus[30]	21	NR	Videos	0	0	0	0	21 (5525)	21 (5525)	0	0

BBPS, Boston Bowel Preparation Scale[59]; *NR*, not reported; *WL*, white-light endoscopy.
*Complete database contains 112 patients; however, data only available for 2 patients.

Data extraction

Details from the included databases were extracted in duplicate using a prespecified data extraction form, including the items of the metadata checklist (Supplementary Tables 1 and 2, available online at www.giejournal.org). This extraction form was piloted on the first 10 databases (B.B.S.L.H. and K.J.N.).

RESULTS

Databases identified by literature and targeted search

The flowchart of the study and database selection process is shown in Figure 1. Of 3524 citations identified by the literature and targeted search, 354 were assessed to be eligible for full-text review. From these articles, 40 unique colonoscopic imaging databases were identified and assessed for further review.

Accessibility databases

Of the 40 unique databases, 22 were open access, from which raw data could be directly downloaded[12-35]; 3 databases were open access with barriers, from which data could not be downloaded[36-38]; and 15 databases had regulated data access.[2,3,39-51] Details and applications of the identified open access database category are described in Tables 1 to 4 and Figure 2. Details of the databases in the open access with barriers and regulated access categories are shown in Supplementary Table 3 (available online at www.giejournal.org).

From the 22 accessible databases, 19 databases focused on polyp detection, 16 on polyp localization, 11 on polyp segmentation, 6 on polyp characterization and 3 on quality

TABLE 3. Continued

Endoscopy system brand	Imaging modality	Image/frame resolution (pixel)s	File format	Histopathology, class, and distribution	Additional polyp characterization information
NR	WL	NR	TIF	613 adenomas; 288 non-adenomas	No
Olympus	NBI, WL	768 × 576	MP4	40 adenomas; 15 serrated adenomas; 21 HPs	Diagnosis endoscopist
NR	WL	224 × 224	JPG	—	38 adenomas; 38 HPs assigned by consensus 3 endoscopists
Olympus	NBI,WL	854 × 480, 1920 × 1080	PNG	8 invasive carcinomas; 12 HGDs; 1 low-grade dysplasia; 8 HPs, 41 nondysplastic lesions; 1 unknown	Size, morphology,† diagnosis endoscopist (NICE classification‡)
Olympus	NBI	1126 × 900, 1920 × 1080	PNG	11 colorectal cancers; 1450 adenomas; 197 sessile serrated lesions; 327 HPs; 5 TSAs; 26 normal mucosa; 38 missing; 15 other	Size, morphology,† location, diagnosis endoscopist, confidence level
Olympus	WL	416 × 416	JPEG	1 invasive carcinoma: 4 high-grade adenomas: 2 TSAs; 4 sessile serrated lesion s; 82 low-grade adenomas: 7 HPs	Size, morphology,* location

TABLE 4. Continued

Endoscopy system brand	Imaging modality	Image, frame resolution (pixels)	File format	Source of ground truth annotation: Classes and distribution	Source of ground truth annotation: Annotation performed by
Olympus	WL	850 × 750	JPG, PNG	Insertion, withdrawal, turning point (ie, cecum) (no. of frames NR)	Expert physician(s)
Olympus, Pentax	NR	332 × 352 to 1921 × 1073	JPEG, AVI	BBPS 0-1: 646 images, 34 videos; BBPS 0-2: 1148 images, 31 videos; cecum (ileocecal valve or appendiceal orifice): 1009 images and 9 videos; ileum 9 images; rectum in retroflexion: 391 images	Combination junior physician and expert physician
Olympus	WL	720 × 576	JPEG	BBPS 0: 500 frames; BBPS 1: 2700 frames, BBPS 2: 975 frames; BBPS 3: 1350 frames	Expert physician(s)

of colonoscopy. Of the available databases, we were able to access 19,463 images, 1,204,772 video frames, and 952 videos from at least 2096 patients. Although the number of video frames registered in these databases was generally high, the median number of unique patients per database was 42 (Q1-Q3 range, 21-300) and the median number of unique polyps per database was 73.5 (Q1-Q3 range, 23.5-132) (Supplementary Figure 1 available online at www.giejournal.org). Six of 22 accessible databases reported on more than 100 different polyps and patients, and only 1 database included more than 100 different polyps and patients (POLAR database,[32] with 802 patients and 2069 polyps). Where reported, in most databases the database acquisition was prospective (14/15; 93%). Of the 22 databases (73%), 6 databases (73%) were from a single center, 3 databases (14%) from multiple centers, and 2 databases (9%) did not report on the number of centers collecting data. Eleven databases were used multiple times by other research groups to train or test their models, whereas 11 were used only once or not at all (range of use for training, 0-165; range of use for testing, 0-19) (Supplementary Table 4). The number of times open access databases were used by other research groups increased significantly in recent years Supplementary Figure 2.

Of the 19 databases focusing on polyp detection, localization, and/or segmentation, the annotation of polyps was performed by using a file name (n = 3), a rectangular bounding box enclosing the polyp (n = 5), a mask (ie, pixel-wise binary mask or polygon-shaped mask) (n = 7), or a combination of the previously mentioned methods (n = 4). Annotation of data was performed by a combination of nonexpert and expert physicians (n = 5), combination of a junior physician and expert (n = 2), expert physicians (n = 5), physicians with unspecified experience (n = 1), and an annotation tool (n = 1). For 6 databases (5/19; 26%) it was not clear who performed the annotation. For purposes of polyp characterization, 6

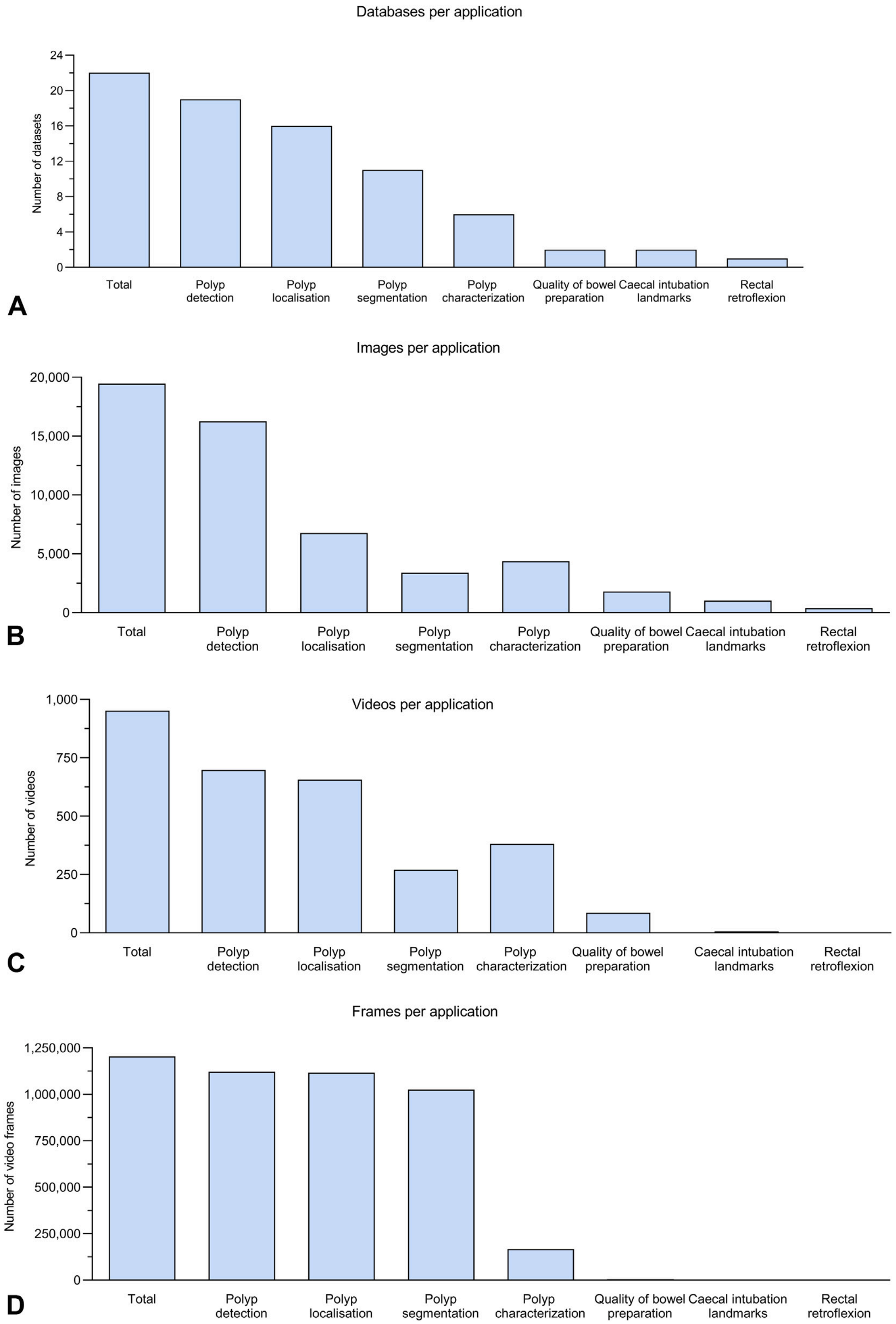


Figure 2. A, Databases per application. **B,** Images per application. **C,** Videos per application. **D,** Frames per application.

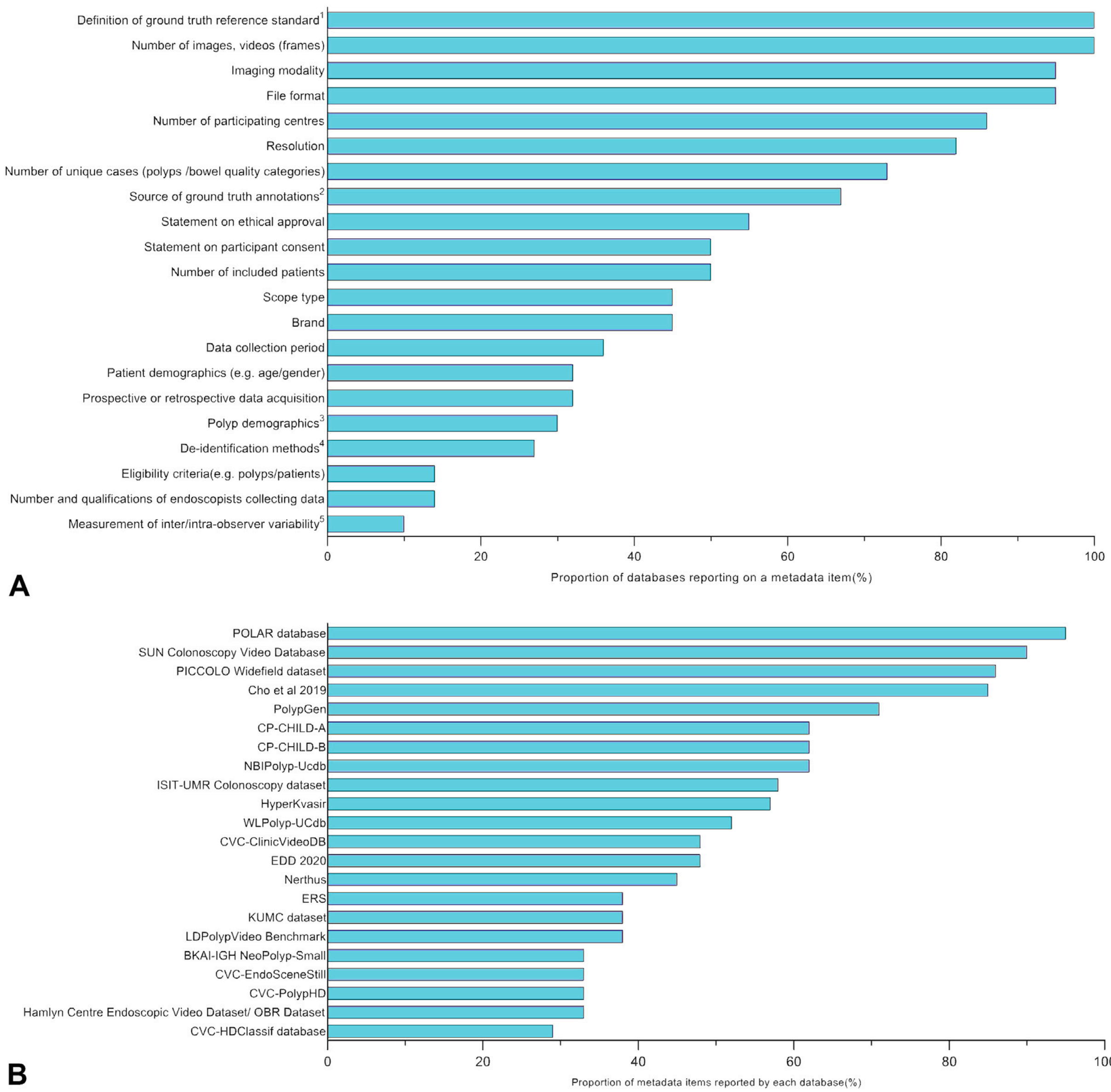


Figure 3. A, Proportion of databases reporting on each metadata item. **B,** Proportion of metadata items reported by each database (**B**). [1]Included detailed, specific definitions of the ground truth annotation, sufficient to allow replications. [2]Source of ground truth annotations, qualifications, and preparation of annotators. Only applicable to polyp detection, polyp localization, polyp segmentation, and/or quality of colonoscopy databases. [3]Describing methods by which data have been deidentified and how protected health information has been removed according to U.S. (Health Insurance Portability and Accountability Act), European (General Data Protection Regulation), or relevant laws. [4]Measurement of inter- and intraobserver variability; methods to mitigate variability and/or resolve discrepancies. Only applicable to polyp detection, polyp localization, polyp segmentation, and/or quality of colonoscopy databases.

publicly available databases were available. The histopathologic annotation categories used were different between all 6 databases (number of annotation categories ranged from 2 to 8 categories). Some databases provided in addition to the histopathology outcome the optical diagnosis of the endoscopist, polyp size, polyp location, and/or polyp morphology. The 3 databases that reported on polyp morphology did so according to the Paris classification[52]; of all polyps in these databases, the proportion of 0-I (I_p+I_s) polyps ranged between 49% and 81%, whereas the proportion of 0-II (II_a+II_b) polyps ranged between 19% and 51% (only 2 II_b cases). Of the 3 publicly available databases focusing on quality of colonoscopy, the Nerthus database[30] focused purely on the quality of the bowel preparation; the Cho database[13] on insertion, withdrawal, and turning point (ie, cecum); and the Hyperkvasir database[22] quality of bowel preparation, cecal intubation landmarks, and the rectum in retroflexion.

TABLE 5. Metadata reporting (for each data item Y/N)

Database	1. Statement on participant consent	2. Statement on ethical approval	3. Data collection period	4. Prospective or retrospective data acquisition	5. Number of participating centers	6. Number and qualifi cations of endos copists collecting data	7. Eligibility criteria (eg, polyps/ patients)	8. De-identification methods[1]	9. Definition of ground truth reference standard[2]	10. Source of ground truth annota tions[3]
BKAI-IGH NeoPolyp-Small[12]	No	No	No	No	No	No	No	No	Yes	No
Cho et al 2019[13]	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes
CP-CHILD-A[14]	Yes	Yes	Yes	No	Yes	No	No	No	Yes	Yes
CP-CHILD-B[14]	Yes	Yes	Yes	No	Yes	No	No	No	Yes	Yes
CVC-ClinicVideoDB[15,16]	No	No	No	No	Yes	No	No	No	Yes	No
CVC-EndoSceneStill[17]	No	No	No	No	Yes	No	No	No	Yes	No
CVC-HDClassif database[18,19]	No	No	No	No	Yes	No	No	No	Yes	No
CVC-PolypHD[17]	No	No	No	No	Yes	No	No	No	Yes	No
EDD 2020[20]	Yes	Yes	No	No	Yes	No	No	Yes	Yes	Yes
ERS[35]	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Hamlyn Centre Endoscopic Video Dataset[21]	No	No	No	No	No	No	No	No	Yes	Yes
HyperKvasir[22]	Yes	Yes	Yes	No	Yes	No	No	Yes	Yes	Yes
ISIT-UMR Colonoscopy dataset[23]	No	Yes	No	Yes	No	No	No	No	Yes	NA
KUMC dataset[24-26]	No	No	No	No	Yes	No	No	No	Yes	Yes
LDPolypVideo Benchmark[27]	No	No	No	No	Yes	No	No	Yes	Yes	No
NBIPolyp-Ucdb[28]	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
WLPolyp-UCdb[29]	Yes	Yes	No	No	Yes	No	No	No	Yes	No
Nerthus[30]	No	No	No	No	Yes	No	No	No	Yes	Yes
PICCOLO Widefield dataset[31]	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
POLAR database [32]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PolypGen[33]	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes
SUN Colonoscopy video database[34]	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Proportion of databases reporting on these data items, %	50	55	36	32	86	14	14	27	100	67

NA, Not applicable.

[1]Describing methods by which data have been deidentified and how protected health information has been removed according to U.S. (Health Insurance Portability and Accountability Act), European (General Data Protection Regulation), or relevant laws.

[2]Included detailed, specific definitions of the ground truth annotation and classes, sufficient to allow replications.

[3]Source of ground truth annotations, qualifications, and preparation of annotators. Only applicable to polyp detection, polyp localization, polyp segmentation, and/or quality of colonoscopy databases.

[4]Measurement of inter- and intraobserver variability; methods to mitigate variability and/or resolve discrepancies. Only applicable to polyp detection, polyp localization, polyp segmentation, and/or quality of colonoscopy databases.

TABLE 5. Continued

11. Measurement of inter/intra-observer variability [4]	12. Patient demographics (eg, age/gender)	13. Polyp demographics (eg, distribution polyp histology type, size, and morphology)	14. Number of included patients	15. Number of unique cases (polyps/bowel quality categories)	16. Number of images, videos (frames)	17. Imaging modality	18. Brand endoscopy system	19. Scope type	20. File format	21. Resolution	Proportion of metadata items reported by this database (%)
No	No	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	33
No	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	85
No	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	62
No	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	62
No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	48
No	No	No	Yes	No	Yes	Yes	No	No	Yes	Yes	33
No	No	Yes	No	No	Yes	Yes	No	No	Yes	No	29
No	No	No	No	Yes	Yes	Yes	No	No	Yes	Yes	33
No	No	No	No	No	Yes	Yes	No	Yes	Yes	No	48
No	No	No	Yes	Yes	Yes	No	No	No	Yes	No	38
No	No	No	No	Yes	Yes	Yes	No	No	Yes	Yes	33
Yes	No	No	No	No	Yes	Yes	No	No	Yes	Yes	57
NA	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	58
No	No	No	No	Yes	yes	Yes	No	No	Yes	Yes	38
No	No	No	No	Yes	Yes	Yes	No	No	Yes	Yes	38
No	Yes	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	62
No	Yes	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	52
No	No	NA	Yes	Yes	Yes	Yes	No	No	Yes	Yes	45
No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	86
No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	95
No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	71
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	90
10	32	30	50	73	100	95	45	45	95	82	

Usability of open access databases

The reporting on 21 essential details of data (ie, metadata items) across all open access databases is shown in Figure 3 and Table 5. The proportion of metadata items reported by the included databases ranged from 29% to 95%. The recently published SUN Colonoscopy video database[34] and POLAR database[32] reported on, respectively, 90% and 95% of the included metadata items, whereas the CVC-HDClassif database[15,16] only reported on 29% of them. Although technical details relating to file format, imaging modality, and resolution were in general well reported (reported by >80% of all databases), patient and polyp characteristics and the annotation process were not completely reported across all databases (reported by <40% of all databases).

DISCUSSION

Large, high-quality, annotated colonoscopic imaging databases are urgently needed to develop and validate AI systems and to facilitate comparisons between systems. This review systematically identified the colonoscopic imaging databases that are currently publicly or, under certain conditions, available and reported on the completeness of their metadata. From our search of the scientific literature and online database search engines, we identified 22 unique publicly available colonoscopic imaging databases that contained 19,463 images, 1,204,772 video frames, and 952 videos. Most databases focused on polyp detection, localization, and/or segmentation, whereas only a small number focused on polyp characterization and quality of colonoscopy.

Visibility of the current accessible AI databases appears to be an issue. Although our review identified 22 open access databases, only 11 of these databases have been used multiple times by other researchers to develop, train, or benchmark their AI system (eg, HyperKvasir[22] and CVC-EndoScenestill[17] databases). This difference in use might have led to lost research opportunities and might have resulted in selection bias because of an overuse of a few potentially nonrepresentative databases. In this regard, it would be useful to improve their visibility, for example by establishing a website with links to all databases to be downloaded. Our study provides an initial point of access that will improve their visibility. Further considerations in this regard include a greater clarity around the accessibility of the imaging data of studies. Numerous authors were e-mailed because they did not mention explicitly in their article whether their database was accessible. As for the future, (medical) journals should recommend authors to report on data accessibility

Publicly available databases can be a powerful enabler to development of AI systems, allowing studies to benchmark their performance in a consistent manner. Whether a database is useful for AI system developers is subjective but is closely associated with the quality and representativeness of that database. Important in this respect is that it should be possible to verify these facets of the database. Therefore, to map the potential usability of the included databases as best as possible, this review extracted 21 essential details (ie, metadata) of each database using a self-developed checklist derived from the Checklist for Artificial Intelligence in Medical Imaging. According to this checklist, this review demonstrated that although technical details were in general well reported, clinical details and annotation details, like measurements of interobserver variability, were under-reported in almost all databases. In about half of the databases, it was not clear what type of patients and polyps were involved and whether the data were retrospectively or prospectively collected. This missing information might limit the appropriate use of the available data, because it is not possible to evaluate whether there is appropriate representation of certain patients and polyps within the database.

Data from certain patients and polyp categories might be more difficult to collect because of varieties or rare occurrence in clinical practice. These less frequently present categories might actually be even more relevant for development of AI systems. This is the case, for instance, for flat polyps (types 0-IIb and 0-IIc). Those polyp types were under-represented in the 3 databases reporting on polyp morphology,[31,32,34] whereas AI systems would especially be useful for the detection of these flat polyps. The same might be true for the diminutive hyperplastic polyps in the rectosigmoid. In daily practice, endoscopists usually leave them in situ, and therefore it is likely that these lesions are under-represented in the current databases focusing on polyp characterization. This potentially inappropriate representation of the targeted research on patients or polyps, which cannot be verified because of nonreporting of metadata, limits the ability of researchers to assess the usability of these data. In addition to nonreporting limiting the appropriate use of data, obtaining generalizable data from most of the current databases is also an issue. Although the number of image frames registered in these databases was generally high, almost all databases included less than 100 different polyps and patients.

As stated earlier, databases in this review reported few details regarding the annotation processes. Details about the number of annotators, annotators' expertise, and consensus process used for discrepancies among annotators were missing in approximately 40% of the included databases. Although these might be unimportant from a technical perspective, they are crucial for clinical interpretation. Many assumptions are made during annotation of data. The incomplete reporting on annotation details may again lead to inappropriate use of data, which in turn may lead to biased study results. Therefore, providing more details regarding the annotation process is paramount because it carries implications for any model trained with these annotations. The POLAR database,[32] SUN Colonoscopy video database,[34] and PICCOLO Widefield database[31] are examples of databases (consisting of, respectively, 2069, 100, and 71 unique polyps) in which a detailed description of polyps and patients with detailed annotation is present. They provide for each polyp a bounding box or binary mask, the histopathology outcome, polyp size, morphology, and/or location. Because these databases report on more of these important characteristics, future users can evaluate better whether they can use these databases for their research.

As described, reporting on essential details of data by database providers is necessary for database consumers to be sufficiently well informed, select appropriate databases for their tasks, and avoid unintentional misuse. However, as this review shows, a wide variety in the proportion of metadata items is reported by each database (29%-95%). Clearly, there are challenges associated with providing details of the collected data. With the inclusion of detailed metadata, which increases the chance of traceability of patient information, privacy becomes a concern. Also, the investment of time, skill, and money to collect, annotate, and store data generates substantial value to data. Therefore, it is logical that such a database is unlikely to be freely available. However, privacy and investment concerns should be balanced with the potential harm implicated by widespread use of potentially unrepresentative data. In addition, it is difficult to create a high-quality, diverse database on your own, in particular for conditions with a low prevalence like early colorectal cancer. Therefore, medical professionals and technicians should set up an online collaborative registry to collect a high-quality and diverse database. To use these data for AI system development, people can pay a small fee, innclude agreement, or collect data for the registry. The risk of traceability of data in this registry could be minimized by compliance with widely adopted guidelines for sharing raw clinical trial data.[53,54]

To the best of our knowledge, this is the largest review to curate a comprehensive list of colonoscopic imaging databases that are publicly or under certain conditions available. A key feature of this review is the broad search strategy to include scientific and online search engines, including those specifically targeting databases. This method proves that not all relevant databases are identified using scientific search engines alone.

We recognize there are several limitations to our study. The searches were run in February 2022. Because of the rapid influx of AI research studies and the probable delays between studies being completed and becoming visible online, we may have missed the latest available databases. Second, because our aim was specifically targeted toward identifying AI-related studies describing colonoscopic imaging databases, we might have missed databases in non–AI-related studies (eg, stored images in a detection study comparing 2 imaging techniques). We limited our search to this type of study to maximize the chance of appropriately annotated data. However, colonoscopic imaging data available in non–AI-related imaging studies might be of value as well. Last, some databases with open access might not have been identified if no statements were made regarding availability. Where status of availability was unclear, we took reasonable measures to access their source. We e-mailed the contact person of the database and allowed a 2-week period for response. Still, we may have missed open access databases, as often we did not receive a response (79%; 88/111 e-mails) or were unable to identify a correct e-mail address (5%; 6/111). However, we did not receive any additional responses in the 2 weeks after the 2-week response time deadline.

In conclusion, this review provides greater insight into the publicly available colonoscopic imaging databases as resources for AI research. The incomplete reporting of metadata of the current databases limits the appropriate use of accessible data, because it is important to be able to evaluate whether there is appropriate annotation and representation of certain patients and polyps within the database. Improved visibility and reporting of metadata of these and future databases would enable researchers to use appropriate databases for their needs. This potentially enhances the quality of future AI research in colonoscopy and ultimately the quality of colonoscopy in daily practice.

ACKNOWLEDGMENTS

We are grateful for the support of F. S. van Etten Jamaludin, librarian with the University of Amsterdam, who assisted with the systematic literature search for this review. We also thank Djuna de Jong for her diligent proofreading of the manuscript.

REFERENCES

1. Misawa M, Kudo SE, Mori Y, et al. Current status and future perspective on artificial intelligence for lower endoscopy. Dig Endosc 2021;33: 273-84.
2. Wang P, Berzin TM, Glissen Brown JR, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut 2019;68:1813-9.
3. Wang P, Liu X, Berzin TM, et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. Lancet Gastroenterol Hepatol 2020;5:343-51.
4. Mori Y, Kudo SE, Misawa M, et al. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. Ann Intern Med 2018;169:357-66.
5. Kaminski MF, Wieszczy P, Rupinski M, et al. Increased rate of adenoma detection associates with reduced risk of colorectal cancer and death. Gastroenterology 2017;153:98-105.
6. Vleugels JLA, Koens L, Dijkgraaf MGW, et al. Suboptimal endoscopic cancer recognition in colorectal lesions in a national bowel screening programme. Gut 2020;69:977-80.
7. Hewett DG, Kaltenbach T, Sano Y, et al. Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. Gastroenterology 2012;143:599-607.
8. Deng J, Dong W, Socher R, et al. Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009, 248-55.
9. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372: n71.
10. Khan SM, Liu X, Nath S, et al. A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability. Lancet Digital Health 2021;3:e51–e66.
11. Mongan J, Moy L, Kahn CE Jr. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers. Radiol Artif Intell 2020;2:e200029.
12. Lan PN, An NS, Hang DV, et al. NeoUNet: towards accurate colon polyp segmentation and neoplasm detection. arXiv preprint arXiv:2107.05 023 2021.

13. Cho M, Kim JH, Hong KS, et al. Identification of cecum time-location in a colonoscopy video by deep learning analysis of colonoscope movement. PeerJ 2019;2019:e7256.
14. Wang W, Tian J, Zhang C, et al. An improved deep learning approach and its applications on colonic polyp images detection. BMC Med Imag 2020;20:83.
15. Angermann Q, Bernal J, Sanchez-Montes C, et al. Real-time polyp detection in colonoscopy videos: a preliminary study for adapting still frame-based methodology to video sequences analysis. Int J Comput Assist Radiol Surg 2017;12:S20-1.
16. Bernal J, Histace A, Masana M, et al. Polyp detection benchmark in colonoscopy videos using GTCreator: a novel fully configurable tool for easy and fast annotation of image databases. Int J Comput Assist Radiol Surg 2018;13:S166-7.
17. Vázquez D, Bernal J, Sánchez FJ, et al. A benchmark for endoluminal scene segmentation of colonoscopy images. J Healthc Eng 2017;2017:4037190.
18. Bernal J, Histace A. GTCreator: a flexible annotation tool for image-based datasets. 2019;14:191-201.
19. Sanchez-Montes C, Sanchez FJ, Bernal J, et al. Computer-aided prediction of polyp histology on white light colonoscopy using surface pattern analysis. Endoscopy 2019;51:261-5.
20. Ali S, Dmitrieva M, Ghatwary N, et al. Deep learning for detection and segmentation of artefact and disease instances in gastrointestinal endoscopy. Med Image Anal 2021;70:102002.
21. Ye M, Giannarou S, Meining A, et al. Online tracking and retargeting with applications to optical biopsy in gastrointestinal endoscopic examinations. Med Image Anal 2016;30:144-57.
22. Borgli H, Thambawita V, Smedsrud PH, et al. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. Sci Data 2020;7:283.
23. Mesejo P, Pizarro D, Abergel A, et al. Computer-aided classification of gastrointestinal lesions in regular colonoscopy. IEEE Trans Med Imag 2016;35:2051-63.
24. Patel K, Li K, Tao K, et al. A comparative study on polyp classification using convolutional neural networks. PloS One 2020;15:e0236452.
25. Li K, Fathan MI, Patel K, et al. Colonoscopy polyp detection and classification: dataset creation and comparative evaluations. PloS One 2021;16:e0255809.
26. Wang G. replication data for: colonoscopy polyp detection and classification: dataset creation and comparative evaluations. V1 ed. Harvard Dataverse; 2021. Available at: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FCBUOR. Accessed October 24, 2022.
27. Ma Y, Chen X, Cheng K, et al. LDPolypVideo benchmark: a large-scale colonoscopy video dataset of diverse polyps. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021, pp 387-96.
28. Figueiredo IN, Pinto L, Figueiredo PN, et al. Unsupervised segmentation of colonic polyps in narrow-band imaging data based on manifold representation of images and Wasserstein distance. Biomed Signal Process Control 2019;53:1-8.
29. Figueiredo PN, Figueiredo IN, Pinto L, et al. Polyp detection with computer-aided diagnosis in white light colonoscopy: comparison of three different methods. Endosc Int Open 2019;7:E209-15.
30. Pogorelov K, Randel KR, de Lange T, et al. Nerthus: a bowel preparation quality video dataset. In: Proceedings of the 8th ACM on Multimedia Systems Conference.
31. Sánchez-Peralta LF, Pagador JB, Picón A, et al. PICCOLO white-light and narrow-band imaging colonoscopic dataset: a performance comparative of models and datasets. Appl Sci 2020;10:8501.
32. Houwen BBSL, Giotis I, Vleugels JLA, et al, on behalf of the POLAR study group. Development and validation of a computer-aided diagnosis system for characterization of diminutive colorectal polyps during live endoscopy: a multicentre study with benchmarking against screening endoscopists. Endoscopy 2022;54:eP133.
33. Ali S, Jha D, Ghatwary N, et al. PolypGen: a multi-center polyp detection and segmentation dataset for generalisability assessment. arXiv preprint arXiv:2106.04463 2021.
34. Misawa M, Kudo SE, Mori Y, et al. Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video). Gastrointest Endosc 2021;93:960-7.e3.
35. Cychnerski J, Dziubich T, Brzeski A. ERS: a novel comprehensive endoscopy image dataset for machine learning, compliant with the MST 3.0 specification. arXiv preprint arXiv:2201.08746 2022.
36. Du N, Wang X, Guo J, et al. Attraction propagation: a user-friendly interactive approach for polyp segmentation in colonoscopy images. PloS One 2016;11:e0155371.
37. Krenzer A, Banck M, Makowski K, et al. A real-time polyp detection system with clinical application in colonoscopy using deep convolutional neural networks. 2022. Available at: https://www.semanticscholar.org/paper/A-Real-Time-Polyp-Detection-System-with-Clinical-in-Krenzer-Banck/a7a06822be037295b7bbc3d63b795e9942fc24d6#extracted. Accessed October 24, 2022.
38. Nogueira-Rodríguez A, Domínguez-Carbajales R, Campos-Tato F, et al. Real-time polyp detection model using convolutional neural networks. Neural Comput Appl 2022;34:10375-96.
39. Poon CCY, Jiang Y, Zhang R, et al. AI-doscopist: a real-time deep-learning-based algorithm for localising polyps in colonoscopy videos with edge computing devices. npj Digit Med 2020;3:1-8.
40. Gong D, Wu L, Zhang J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. Lancet Gastroenterol Hepatol 2020;5:352-61.
41. Lee JY, Jeong J, Song EM, et al. Real-time detection of colon polyps during colonoscopy using deep learning: systematic validation with four independent datasets. Sci Rep 2020;10:8379.
42. Lee JY, Calderwood AH, Karnes W, et al. Artificial intelligence for assessment of bowel preparation. Gastrointest Endosc 2022;95:512-18. e1.
43. Low DJ, Hong Z, Khan R, et al. Automated detection of cecal intubation with variable bowel preparation using a deep convolutional neural network. Endosc Int Open 2021;9:E1778-84.
44. Lui TKL, Wong KKY, Mak LLY, et al. Endoscopic prediction of deeply submucosal invasive carcinoma with use of artificial intelligence. Endosc Int Open 2019;7:E514-20.
45. Renner J, Phlipsen H, Haller B, et al. Optical classification of neoplastic colorectal polyps-a computer-assisted approach (the COACH study). Scand J Gastroenterol 2018;53:1100-6.
46. Repici A, Badalamenti M, Maselli R, et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. Gastroenterology 2020;159:512-20.
47. Saito H, Tanimoto T, Ozawa T, et al. Automatic anatomical classification of colonoscopic images using deep convolutional neural networks. Gastroenterol Rep 2021;9:226-33.
48. Sinonquel P, Eelbode T, Hassan C, et al. Real-time unblinding for validation of a new CADe tool for colorectal polyp detection. Gut 2021;70:641.
49. Urban G, Tripathi P, Alkayali T, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology 2018;155:1069-78.
50. Wang P, Xiao X, Glissen Brown JR, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng 2018;2:741-8.
51. Yamada M, Saito Y, Imaoka H, et al. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. Sci Rep 2019;9:14465.
52. Participants in the Paris Workshop. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. Gastrointest Endosc 2003;58:S3-43.
53. Hrynaszkiewicz I, Norton ML, Vickers AJ, et al. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers. BMJ 2010;340:c181.
54. Keerie C, Tuck C, Milne G, et al. Data sharing in clinical trials—practical guidance on anonymising trial datasets. Trials 2018;19:25.
55. Bernal J, Sánchez J, Vilarino F. Towards automatic polyp detection with a polyp appearance model. Pattern Recogn 2012;45:3166-82.

56. Bernal J, Sánchez FJ, Fernández-Esparrach G, et al. WM-DOVA maps for accurate polyp highlighting in colonoscopy: validation vs. saliency maps from physicians. Comput Med Imaging Graph 2015;43:99-111.
57. Pogorelov K, Randel KR, Griwodz C, et al. Kvasir: a multi-class image dataset for computer aided gastrointestinal disease detection. In: Proceedings of the 8th ACM on Multimedia Systems Conference.
58. Jha D, Smedsrud PH, Riegler MA, et al. Kvasir-seg: a segmented polyp dataset. In: International Conference on Multimedia Modeling: Springer; 2020, pp 451-62.
59. Lai EJ, Calderwood AH, Doros G, et al. The Boston Bowel Preparation Scale: a valid and reliable instrument for colonoscopy-oriented research. Gastrointest Endosc 2009;69:620-5.

Abbreviation: AI, artificial intelligence.

DISCLOSURE: *The following authors disclosed financial relationships: P. Fockens: Research support from Boston Scientific; consultant for Olympus and Cook Endoscopy. E. Dekker: Research support from Fujifilm; consultant for Olympus, Fujifilm, GI Supply, PAION, Ambu, and CPP-FAP; speaker for Olympus, Roche, GI Supply, Norgine, Fujifilm, and IPSEN. All other authors disclosed no financial relationships.*

0016-5107
https://doi.org/10.1016/j.gie.2022.08.043

Received June 14, 2022. Accepted August 30, 2022.

Current affiliations: Department of Gastroenterology and Hepatology, Amsterdam Gastroenterology Endocrinology Metabolism, Amsterdam University Medical Centres, location Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands (1), Department of Gastroenterology and Hepatology, Radboud University Nijmegen Medical Center, Radboud University of Nijmegen, Nijmegen, the Netherlands (2).

Reprint requests: Evelien Dekker, MD, PhD, Amsterdam University Medical Center, location Academic Medical Center, Department of Gastroenterology and Hepatology, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.

APPENDIX

APPENDIX 1. DATABASE DEFINITIONS

A database was classified as a polyp detection dataset if the ground truth consisted of file names or file (eg, Excel file) that indicated whether there was a polyp visible on each image or video. In case of a polyp localization database, the ground truth had to consist of a specification of the area in which a polyp was visible within an image or video. This could have be done, for example, by means of a bounding box, which is a 2-dimensional rectangle enclosing the polyp. To classify a database as a polyp segmentation database, for the ground truth a binary mask had to be created of the polyp within the image or video (ie, pixel-wise mask). A database was classified as a characterization database if the ground truth consisted of annotated characterization categories (eg, for polyp characterization, the histologic subtype diagnosed by a pathologist for bowel preparation score, eg, the Boston Bowel Preparation Scale [see reference 1 in the Supplementary References] scored by an endoscopist).

APPENDIX 2. SEARCH STRATEGY

The search performed by Faridi van Etten-Jamaludin (Amsterdam UMC, Medical Library AMC) on February 14, 2022.

PubMed, Embase (Ovid)	Before deduplication	After deduplication
Total	4360	3470

PubMed: 2392 hits

("Computing Methodologies"[MAJR] OR neural network*[tiab] OR artificial intelligence[tiab] OR computer aided[tiab] OR computational method*[tiab] OR machine learning[tiab]) AND ("Diagnostic Imaging"[Mesh] OR "diagnostic imaging" [Subheading] OR "Video Recording"[-Mesh] OR imaging[tiab] OR image*[tiab] OR video*[tiab]) AND ("Colonoscopy"[Mesh] OR "Colonic Polyps"[Majr] OR "Colorectal Neoplasms"[Majr] OR colonoscop*[tiab] OR endoscop*[tiab] OR bowel preparation*[tiab] OR cecal intubation[tiab] OR caecal intubation[tiab]) AND ("Algorithms"[Mesh] OR dataset*[tiab] OR database*[tiab] OR data file*[tiab] OR data collection*[tiab] OR real-time image recognition system*[tiab] OR textural surface pattern*[tiab] OR algorithm*[tiab] OR frame*[tiab] OR bowel preparation*[tiab])

EMBASE (OVID): 1968 hits

No.	Searches	Results
1	exp machine learning/ or artificial intelligence/ or computer analysis/ or (neural network* or artificial intelligence or computer aided or computational method* or machine learning).ti,ab,kw.	503,749
2	diagnostic imaging/ or computer assisted diagnosis/ or (imaging or image* or video*).ti,ab,kw.	2,127,922
3	exp colonoscopy/ or exp colon polyp/ or colorectal polyp/ or exp colorectal tumor/ or (colonoscop* or endoscop* or bowel preparation* or cecal intubation or caecal intubation).ti,ab,kw.	477,944
4	exp algorithm/ or image analysis/ or (dataset* or database* or data file* or real-time or pattern* or algorithm*).ti,ab,kw.	3,667,637
5	(data adj3 (set* or collect*)).ti,ab,kw.	739,279
6	4 or 5	4,241,788
7	1 and 2 and 3 and 6	1968

Google search engine, a Google Dataset Search: 296 hits

Search performed by Drs Houwen and Nass on February 14, 2022.

(computer or computing or artificial intelligence or neural networks) AND (image or imaging or video* or vision*) AND (dataset* or database* or data file* or data collection*) AND (polyp* or endoscopy or colonoscopy or colorectal)

Github and Figshare:

Search performed by Drs Houwen and Nass on February 14, 2022.

SUPPLEMENTARY REFERENCES

1. Calderwood AH, Jacobson BC. Comprehensive validation of the Boston Bowel Preparation Scale. Gastrointest Endosc 2010;72:686-92.
2. Poon CCY, Jiang Y, Zhang R, et al. AI-doscopist: a real-time deep-learning-based algorithm for localising polyps in colonoscopy videos with edge computing devices. npj Digit Med 2020;3:1-8.
3. Gong D, Wu L, Zhang J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. Lancet Gastroenterol Hepatol 2020;5:352-61.
4. Krenzer A, Banck M, Makowski K, et al. A real-time polyp detection system with clinical application in colonoscopy using deep convolutional neural networks. 2022. Available at: https://www.researchsquare.com/article/rs-1310139/v1.pdf. Accessed October 24, 2022.
5. Lee JY, Jeong J, Song EM, et al. Real-time detection of colon polyps during colonoscopy using deep learning: systematic validation with four independent datasets. Sci Rep 2020;10:8379.

6. Lee JY, Calderwood AH, Karnes W, et al. Artificial intelligence for assessment of bowel preparation. Gastrointest Endosc 2022;95:512-8.e1.
7. Low DJ, Hong Z, Khan R, et al. Automated detection of cecal intubation with variable bowel preparation using a deep convolutional neural network. Endosc Int Open 2021;9:E1778-84.
8. Lui TKL, Wong KKY, Mak LLY, et al. Endoscopic prediction of deeply submucosal invasive carcinoma with use of artificial intelligence. Endosc Int Open 2019;7:E514-20.
9. Du N, Wang X, Guo J, et al. Attraction propagation: a user-friendly interactive approach for polyp segmentation in colonoscopy images. PloS One 2016;11:e0155371.
10. Nogueira-Rodríguez A, Domínguez-Carbajales R, Campos-Tato F, et al. Real-time polyp detection model using convolutional neural networks. Neural Comput Appl 2022;34:10375-96.
11. Renner J, Phlipsen H, Haller B, et al. Optical classification of neoplastic colorectal polyps-a computer-assisted approach (the COACH study). Scand J Gastroenterol 2018;53:1100-6.
12. Repici A, Badalamenti M, Maselli R, et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. Gastroenterology 2020;159:512-20.
13. Saito H, Tanimoto T, Ozawa T, et al. Automatic anatomical classification of colonoscopic images using deep convolutional neural networks. Gastroenterol Rep 2021;9:226-33.
14. Sinonquel P, Eelbode T, Hassan C, et al. Real-time unblinding for validation of a new CADe tool for colorectal polyp detection. Gut 2021;70:641.
15. Urban G, Tripathi P, Alkayali T, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology 2018;155:1069-78.
16. Wang P, Xiao X, Glissen Brown JR, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng 2018;2:741-8.
17. Wang P, Berzin TM, Glissen Brown JR, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut 2019;68:1813-9.
18. Wang P, Liu X, Berzin TM, et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. Lancet Gastroenterol Hepatol 2020;5:343-51.
19. Yamada M, Saito Y, Imaoka H, et al. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. Sci Rep 2019;9:14465.
20. Lan PN, An NS, Hang DV, et al. NeoUNet: towards accurate colon polyp segmentation and neoplasm detection. arXiv preprint arXiv:2107.05023 2021.
21. Cho M, Kim JH, Hong KS, et al. Identification of cecum time-location in a colonoscopy video by deep learning analysis of colonoscope movement. PeerJ 2019;2019:e7256.
22. Wang W, Tian J, Zhang C, et al. An improved deep learning approach and its applications on colonic polyp images detection. BMC Med Imag 2020;20:1-14.
23. Zhao X, Jiang W, Yu T, et al. Artificial intelligence: future for polyp diagnosis! 2021. Available at: https://jasonyanglu.github.io/files/lecture_notes/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0_2021/Project/Artificial%20intelligence_Future%20for%20Polyp%20Diagnosis!.pdf. Accessed October 24, 2022.
24. Angermann Q, Bernal J, Sanchez-Montes C, et al. Real-time polyp detection in colonoscopy videos: a preliminary study for adapting still frame-based methodology to video sequences analysis. Int J Comput Assist Radiol Surg 2017;12:S20-1.
25. Bernal J, Histace A, Masana M, et al. Polyp detection benchmark in colonoscopy videos using GTCreator: a novel fully configurable tool for easy and fast annotation of image databases. Int J Comput Assist Radiol Surg 2018;13:S166-7.
26. Angermann Q, Bernal J, Sánchez-Montes C, et al. Towards real-time polyp detection in colonoscopy videos: adapting still frame-based methodologies for video sequences analysis. In: Computer assisted and robotic endoscopy and clinical image-based procedures. Cham: Springer; 2017. pp 29-41.
27. Angermann Q, Histace A, Hammami M, et al. Hardware platforms benchmark for real-time polyp detection. In: 2017 Euromicro Conference on Digital System Design.
28. Dijkstra W, Sobiecki A, Bernal J, et al. Towards a single solution for polyp detection, localization and segmentation in colonoscopy images. In: VISIGRAPP (4: VISAPP). 2019.
29. Ma Y, Chen X, Sun B. Polyp detection in colonoscopy videos by bootstrapping via temporal consistency. In: 2020 IEEE 17th International Symposium on Biomedical Imaging.
30. Pappalardo G, Farinella GM. On the detection of colorectal polyps with hierarchical fine-tuning. In: 2020 IEEE International Symposium on Medical Measurements and Applications.
31. Shin Y, Qadir HA, Balasingham I. Abnormal colon polyp image synthesis using conditional adversarial networks for improved detection performance. IEEE Access 2018;6:56007-17.
32. Sun X, Wang D, Zhang C, et al. Colorectal polyp detection in real-world scenario: design and experiment study. In: 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence.
33. Li X, Liu R, Li M, et al. Real-time polyp detection for colonoscopy video on CPU. In: 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence.
34. Sun X, Zhang P, Wang D, et al. Colorectal polyp segmentation by u-net with dilation convolution. In: 2019 18th IEEE International Conference on Machine Learning and Applications.
35. Thambawita V, Jha D, Hammer HL, et al. An extensive study on cross-dataset bias and evaluation metrics interpretation for machine learning applied to gastrointestinal tract abnormality classification. ACM Trans Comput Healthc 2020;1:1-29.
36. Zhang P, Sun X, Wang D, et al. An efficient spatial-temporal polyp detection framework for colonoscopy video. In: 2019 IEEE 31st International Conference on Tools with Artificial Intelligence.
37. Jha D, Smedsrud PH, Johansen D, et al. A comprehensive study on colorectal polyp segmentation with ResUNet++, conditional random field and test-time augmentation. IEEE J Biomed Health Inform 2021;25:2029-40.
38. Wu L, Hu Z, Ji Y, et al. Multi-frame collaboration for effective endoscopic video polyp detection via spatial-temporal feature transformation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer; 2021. p 302-12.
39. Ahmad OF, González-Bueno Puyal J, Brandao P, et al. Performance of artificial intelligence for detection of subtle and advanced colorectal neoplasia. Dig Endosc 2022;34:862-9.
40. Li JW, Chia T, Fock KM, et al. Artificial intelligence and polyp detection in colonoscopy: use of a single neural network to achieve rapid polyp localization for clinical use. J Gastroenterol Hepatol 2021;36:3298-307.
41. Zhang Z, Ma L, Chana Y, et al. A practical polyp detecting model in colonoscopy video by post-processing. In: 2021 14th International Congress on Image and Signal Processing: BioMedical Engineering and Informatics; 2021.
42. de Almeida Thomaz V, Sierra-Franco CA, Raposo AB. Training data enhancements for improving colonic polyp detection using deep convolutional neural networks. Artif Intell Med 2021;111:101988.
43. Qadir HA, Balasingham I, Solhusvik J, et al. Improving automatic polyp detection using CNN by exploiting temporal dependency in colonoscopy video. IEEE J Biomed Health Inform 2019;24:180-93.
44. Shin Y, Qadir HA, Aabakken L, et al. Automatic colon polyp detection using region based deep CNN and post learning approaches. IEEE Access 2018;6:40950-62.
45. Xu J, Zhao R, Yu Y, et al. Real-time automatic polyp detection in colonoscopy using feature enhancement module and spatiotemporal similarity correlation unit. Biomed Signal Process Control 2021;66:102503.
46. Vázquez D, Bernal J, Sánchez FJ, et al. A benchmark for endoluminal scene segmentation of colonoscopy images. J Healthc Eng 2017;2017:4037190.
47. Bernal J, Sánchez J, Vilarino F. Towards automatic polyp detection with a polyp appearance model. Pattern Recogn 2012;45:3166-82.

48. Bernal J, Sánchez FJ, Fernández-Esparrach G, et al. WM-DOVA maps for accurate polyp highlighting in colonoscopy: validation vs. saliency maps from physicians. Comput Med Imag Graph 2015;43:99-111.
49. Akbari M, Mohrekesh M, Nasr-Esfahani E, et al. Polyp segmentation in colonoscopy images using fully convolutional network. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.
50. Angermann Q, Histace A, Romain O. Active learning for real time detection of polyps in videocolonoscopy. Proc Comput Sci 2016;90:182-7.
51. Banik D, Roy K, Bhattacharjee D, et al. Polyp-Net: a multimodel fusion network for polyp segmentation. IEEE Trans Instrument Measur 2020;70:1-12.
52. Bardhi O, Sierra-Sosa D, Garcia-Zapirain B, et al. Automatic colon polyp detection using convolutional encoder-decoder model. In: 2017 IEEE International Symposium on Signal Processing and INFORMATION TECHNOLOGY.
53. Bernal J, Núñez JM, Sánchez FJ, et al. Polyp segmentation method in colonoscopy videos by means of msa-dova energy maps calculation. In: Workshop on Clinical Image-Based Procedures. Cham: Springer; 2014, pp 41-9.
54. Bernal J, Sánchez J, Vilarino F. Impact of image preprocessing methods on polyp localization in colonoscopy frames. In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.
55. Billah M, Waheed S. Minimum redundancy maximum relevance (mRMR) based feature selection from endoscopic images for automatic gastrointestinal polyp detection. Multimedia Tools Appl 2020;79:23633-43.
56. Brandao P, Mazomenos E, Ciuti G, et al. Fully convolutional neural networks for polyp segmentation in colonoscopy. In: Medical Imaging 2017: Computer-Aided Diagnosis: International Society for Optics and Photonics; 2017, p. 101-7.
57. Brandao P, Zisimopoulos O, Mazomenos E, et al. Towards a computed-aided diagnosis system in colonoscopy: automatic polyp segmentation using convolution neural networks. J Med Robot Res 2018;3:1840002.
58. Bravo D, Ruano J, Gómez M, et al. Automatic polyp detection and localization during colonoscopy using convolutional neural networks. In: Medical Imaging 2020: Computer-Aided Diagnosis: International Society for Optics and Photonics; 2020, p. 775-85.
59. Bravo D, Ruano J, Gómez M, et al. Automatic polyp localization by low level superpixel features. In: 15th International Symposium on Medical Information Processing and Analysis: International Society for Optics and Photonics; 2020, p. 360-9.
60. Chen C-W, Fuh C-S. Deep learning for polyp detection in colonoscopy. National Taiwan University, Taiwan. Available at: https://www.csie.ntu.edu.tw/~fuh/personal/DeepLearningforPolypDetectioninColonoscopy.pdf. Accessed October 24, 2022.
61. Cho Y-J, Bae S-H, Yoon K-J. Multi-classifier-based automatic polyp detection in endoscopic images. J Med Biol Eng 2016;36:871-82.
62. de Almeida Thomaz V, Sierra-Franco CA, Raposo AB. Training data enhancements for robust polyp segmentation in colonoscopy images. In: 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems. IEEE, 2019. p. 192-97.
63. Deeba F, Bui FM, Wahid KA. Computer-aided polyp detection based on image enhancement and saliency-based selection. Biomed Signal Process Control 2020;55:101530.
64. Fan D-P, Ji G-P, Zhou T, et al. Pranet: parallel reverse attention network for polyp segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2020, p. 263-73.
65. Fang Y, Zhu D, Yao J, et al. ABC-Net: area-boundary constraint network with dynamical feature selection for colorectal polyp segmentation. IEEE Sens J 2020;21:11799-809.
66. Fathan MI. A comparative study on polyp classification and localization from colonoscopy videos. University of Kansas; 2019. Available at: https://kuscholarworks.ku.edu/handle/1808/29702. Accessed October 24, 2022.
67. Feng R, Lei B, Wang W, et al. SSN: a stair-shape network for real-time polyp segmentation in colonoscopy images. In: 2020 IEEE 17th International Symposium on Biomedical Imaging. IEEE, 2020. p. 225-9.
68. Fernández-Esparrach G, Bernal J, López-Cerón M, et al. Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps. Endoscopy 2016;48:837-42.
69. Bernal J, Tajkbaksh N, Sanchez FJ, et al. Comparative validation of polyp detection methods in video colonoscopy: results from the MICCAI 2015 endoscopic vision challenge. IEEE Trans Med Imag 2017;36:1231-49.
70. Gil D, Sánchez FJ, Fernández-Esparrach G, et al. 3D stable spatio-temporal polyp localization in colonoscopy videos. In: Computer-assisted and robotic endoscopy: Springer; 2015, p. 140-52.
71. Guo X, Yang C, Liu Y, et al. Learn to threshold: thresholdnet with confidence-guided manifold mixup for polyp segmentation. IEEE Trans Med Imag 2020;40:1134-46.
72. Guo X, Zhang N, Guo J, et al. Automated polyp segmentation for colonoscopy images: a method based on convolutional neural networks and ensemble learning. Med Phys 2019;46:5666-76.
73. Guo Y, Bernal J, J Matuszewski B. Polyp segmentation with fully convolutional deep neural networks—extended evaluation study. J Imag 2020;6:69.
74. Guo Y, Matuszewski BJ. Polyp segmentation with fully convolutional deep dilation neural network. In: Annual Conference on Medical Image Understanding and Analysis: Springer; 2019, p. 377-88.
75. Haj-Manouchehri A, Mohammadi HM. Polyp detection using CNNs in colonoscopy video. IET Comput Vision 2020;14:241-7.
76. Huang C-H, Xiao W-T, Chang L-J, et al. Automatic tissue segmentation by deep learning: from colorectal polyps in colonoscopy to abdominal organs in CT exam. In: 2018 IEEE Visual Communications and Image Processing. IEEE, 2018. p. 1-4.
77. Hwang M, Wang D, Kong X-X, et al. An automated detection system for colonoscopy images using a dual encoder-decoder model. Comput Med Imag Graph 2020;84:101763.
78. Jenssen R, Kampffmeyer MC. Uncertainty and interpretability in convolutional neural networks for semantic segmentation of colorectal polyps. Med Image Anal 2020;60:101619.
79. Kang J, Gwak J. Ensemble of instance segmentation models for polyp segmentation in colonoscopy images. IEEE Access 2019;7:26440-7.
80. Jha D, Riegler MA, Johansen D, et al. Doubleu-net: a deep convolutional neural network for medical image segmentation. In: 2020 IEEE 33rd International Symposium on Computer-based Medical Systems. IEEE, 2020. p. 558-64.
81. Jha D, Smedsrud PH, Riegler MA, et al. Resunet++: an advanced architecture for medical image segmentation. In: 2019 IEEE International Symposium on Multimedia. IEEE, 2019. p. 225-55.
82. Jia X, Mai X, Cui Y, et al. Automatic polyp recognition in colonoscopy images using deep learning and two-stage pyramidal feature prediction. IEEE Trans Automat Sci Eng 2020;17:1570-84.
83. Kassani SH, Kassani PH, Wesolowski MJ, et al. Automatic polyp segmentation using convolutional neural networks. arXiv preprint arXiv:2004.10792 2020.
84. Kayser M, Soberanis-Mukul RD, Zvereva A-M, et al. Understanding the effects of artifacts on automated polyp detection and incorporating that knowledge via learning without forgetting. arXiv preprint arXiv:2002.02883 2020.
85. Li Q, Yang G, Chen Z, et al. Colorectal polyp segmentation using a fully convolutional neural network. In: 2017 10th International Congress on Image and Signal Processing: Biomedical Engineering and Informatics; 2017. IEEE, 2017. p. 1-5.
86. Liu M, Jiang J, Wang Z. Colonic polyp detection in endoscopic videos with single shot detection based deep convolutional neural network. IEEE Access 2019;7:75058-66.
87. Mahmood F, Yang Z, Chen R, et al. Polyp segmentation and classification using predicted depth from monocular endoscopy. In: Medical Imaging 2019: Computer-Aided Diagnosis: International Society for Optics and Photonics; 2019. SPIE, 2019. p. 268-72.

88. Mahmud T, Paul B, Fattah SA. PolypSegNet: a modified encoder-decoder architecture for automated polyp segmentation from colonoscopy images. Comput Biol Med 2021;128:104119.
89. Mo X, Tao K, Wang Q, et al. An efficient approach for polyps detection in endoscopic videos based on faster R-CNN. In: 2018 24th International Conference on Pattern Recognition. IEEE, 2018. p. 3929-34.
90. Mostafiz R, Rahman MM, Uddin MS. Gastrointestinal polyp classification through empirical mode decomposition and neural features. SN Appl Sci 2020;2:1-10.
91. Nagy S, Lilik F, Koczy LT. Entropy based fuzzy classification and detection aid for colorectal polyps. In: 2017 IEEE AFRICON. IEEE, 2017. p. 78-82.
92. Nguyen N-Q, Vo DM, Lee S-W. Contour-aware polyp segmentation in colonoscopy images using detailed upsampling encoder-decoder networks. IEEE Access 2020;8:99495-508.
93. Nguyen Q, Lee S-W. Colorectal segmentation using multiple encoder-decoder network in colonoscopy images. In: 2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering 2018, IEEE.
94. Patel K, Li K, Tao K, et al. A comparative study on polyp classification using convolutional neural networks. PloS One 2020;15:e0236452.
95. Pogorelov K, Albisser Z, Ostroukhova O, et al. Opensea: open search based classification tool. In: Proceedings of the 9th ACM Multimedia Systems Conference; 2018.
96. Pogorelov K, Ostroukhova O, Jeppsson M, et al. Deep learning and hand-crafted feature based approaches for polyp detection in medical videos. In: 2018 IEEE 31st International Symposium on Computer-Based Medical Systems 2018, IEEE.
97. Poorneshwaran J, Kumar SS, Ram K, et al. Polyp segmentation using generative adversarial network. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2019, IEEE.
98. Qadir HA, Shin Y, Solhusvik J, et al. Polyp detection and segmentation using mask R-CNN: Does a deeper feature extractor CNN always perform better? In: 2019 13th International Symposium on Medical Information and Communication Technology. 2019, IEEE.
99. Qadir HA, Shin Y, Solhusvik J, et al. Toward real-time polyp detection using fully CNNs for 2D Gaussian shapes prediction. Med Image Anal 2021;68:101897.
100. Qadir HA, Solhusvik J, Bergsland J, et al. A framework with a fully convolutional neural network for semi-automatic colon polyp annotation. IEEE Access 2019;7:169537-47.
101. Qian Z, Lv Y, Lv D, et al. A new approach to polyp detection by preprocessing of images and enhanced faster R-CNN. IEEE Sens J 2020;21:11374-81.
102. Sanchez-Gonzalez A, Garcia-Zapirain B, Sierra-Sosa D, et al. Automatized colon polyp segmentation via contour region analysis. Comput Biol Med 2018;100:152-64.
103. Sánchez-González A, Garcia-Zapirain B, Sierra-Sosa D, et al. Colon polyp segmentation using texture analysis. In: 2018 IEEE International Symposium on Signal Processing and Information Technology. 2018, IEEE.
104. Sánchez-Peralta LF, Picón A, Antequera-Barroso JA, et al. Eigenloss: combined PCA-based loss function for polyp segmentation. Mathematics 2020;8:1316.
105. Sánchez-Peralta LF, Picón A, Sánchez-Margallo FM, et al. Unravelling the effect of data augmentation transformations in polyp segmentation. Int J Comput Assist Radiol Surg 2020;15:1975-88.
106. Shafi A, Rahman MM. Decomposition of color wavelet with higher order statistical texture and convolutional neural network features set based classification of colorectal polyps from video endoscopy. Int J Electric Comput Eng 2020;10:2986.
107. Shin Y, Balasingham I. Automatic polyp frame screening using patch based combined feature and dictionary learning. Comput Med Imag Graph 2018;69:33-42.
108. Sornapudi S, Meng F, Yi S. Region-based automated localization of colonoscopy and wireless capsule endoscopy polyps. Appl Sci 2019;9:2404.
109. Taha B, Dias J, Werghi N. Convolutional neural network as a feature extractor for automatic polyp detection. In: 2017 IEEE International Conference on Image Processing. 2017, IEEE.
110. Tajbakhsh N, Chi C, Gurudu SR, et al. Automatic polyp detection from learned boundaries. In: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI): IEEE; 2014.
111. Tajbakhsh N, Gurudu SR, Liang J. A classification-enhanced vote accumulation scheme for detecting colonic polyps. In: International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging: Springer; 2013.
112. Tajbakhsh N, Gurudu SR, Liang J. Automatic polyp detection using global geometric constraints and local intensity variation patterns. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2014 p. 179-87.
113. Tajbakhsh N, Gurudu SR, Liang J. Automated polyp detection in colonoscopy videos using shape and context information. IEEE Trans Med Imag 2015;35:630-44.
114. Takamatsu M, Hasegawa M. Architecture for accurate polyp segmentation in motion-blurred colonoscopy images. In: 2020 19th IEEE International Conference on Machine Learning and Applications.
115. Tashk A, Herp J, Nadimi E. Fully automatic polyp detection based on a novel U-Net architecture and morphological post-process. In: 2019 International Conference on Control, Artificial Intelligence: Robotics & Optimization; 2019.
116. Thanh NC, Long TQ. Polyp segmentation in colonoscopy images using ensembles of U-Nets with EfficientNet and asymmetric similarity loss function. In: 2020 RIVF International Conference on Computing and Communication Technologies.
117. Viscaino M, Cheein FA. Machine learning for computer-aided polyp detection using wavelets and content-based image. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:961-5.
118. Wang D, Zhang N, Sun X, et al. Afp-net: Realtime anchor-free polyp detection in colonoscopy. In: 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI): IEEE; 2019.
119. Wang S, Cong Y, Zhu H, et al. Multi-scale context-guided deep network for automated lesion segmentation with endoscopy images of gastrointestinal tract. IEEE J Biomed Health Inform 2020;25:514-25.
120. Wang Y, Pomeroy M, Cao W, et al. Differentiation of polyps by clinical colonoscopy via integrated color information, image derivatives and machine learning. In: Medical Imaging 2019: Computer-Aided Diagnosis: International Society for Optics and Photonics; 2019. Available at: https://ui.adsabs.harvard.edu/abs/2019SPIE10950E..2YW/abstract. Accessed October 24, 2022.
121. Wichakam I, Panboonyuen T, Udomcharoenchaikit C, et al. Real-time polyps segmentation for colonoscopy video frames using compressed fully convolutional network. In: International Conference on Multimedia Modeling: Springer; 2018.
122. Wickstrøm K, Kampffmeyer M, Jenssen R. Uncertainty modeling and interpretability in convolutional neural networks for polyp segmentation. In: 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing. Available at: https://ieeexplore.ieee.org/document/8516998. Accessed October 24, 2022.
123. Yu J, Pan H, Yin Q, et al. Fully convolutional densenets for polyp segmentation in colonoscopy. In: 2019 IEEE 35th International Conference on Data Engineering Workshops. Available at: https://ieeexplore.ieee.org/document/8750956. Accessed October 24, 2022.
124. Yuan Y, Li D, Meng MQ-H. Automatic polyp detection via a novel unified bottom-up and top-down saliency approach. IEEE J Biomed Health Inform 2017;22:1250-60.
125. Zhang L, Dolwani S, Ye X. Automated polyp segmentation in colonoscopy frames using fully convolutional neural network and textons. In: Annual Conference on Medical Image Understanding and Analysis: Springer; 2017.
126. Zhang R, Li G, Li Z, et al. Adaptive context selection for polyp segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2020.
127. Zheng Y, Zhang R, Yu R, et al. Localisation of colorectal polyps by convolutional neural network features learnt from white light and narrow band endoscopic images of multiple databases. Annu Int Conf IEEE Eng Med Biol Soc 2018;2018:4142-5.

128. Nasrin S, Alavi J, Viswanathan P. Segmentation of polyps in gastrointestinal tract images. University of Alberta; 2021.
129. Yin Z, Liang K, Ma Z, et al. Duplex contextual relation network for polyp segmentation. arXiv preprint arXiv:2103.06725 2021.
130. Rajinikanth V, Aslam SM, Kadry S, et al. Semi/Fully-Automated Segmentation Of Gastric-Polyp Using Aquila-Optimization-Algorithm Enhanced Images. Comput Mater Contin 2022;70:4087-105.
131. Li K, Fathan MI, Patel K, et al. Colonoscopy polyp detection and classification: dataset creation and comparative evaluations. PloS One 2021;16:e0255809.
132. Wang G. Replication data for: colonoscopy polyp detection and classification: dataset creation and comparative evaluations. V1 ed. Harvard Dataverse; 2021. Available at: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FCBUOR. Accessed October 24, 2022.
133. Hasan M, Hossain MM, Mia S, et al. A combined approach of non-subsampled contourlet transform and convolutional neural network to detect gastrointestinal polyp. Multimedia Tools Appl 2022;81:9949-68.
134. Liu X, Guo X, Liu Y, et al. Consolidated domain adaptive detection and localization framework for cross-device colonoscopic images. Med Image Anal 2021;71:102052.
135. Taş M, Yılmaz B. Super resolution convolutional neural network based pre-processing for automatic polyp detection in colonoscopy images. Comput Electric Eng 2021;90:106959.
136. Ellahyani A, Jaafari IE, Charfi S, et al. Fine-tuned deep neural networks for polyp detection in colonoscopy images. Person Ubiq Comput 2022:1-13.
137. Cheng M, Kong Z, Song G, et al. Learnable oriented-derivative network for polyp segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021.
138. Su H, Lin B, Huang X, et al. MBFFNet: multi-branch feature fusion network for colonoscopy. Front Bioeng Biotechnol 2021;9:1-14.
139. Zhao X, Zhang L, Lu H. Automatic polyp segmentation via multi-scale subtraction network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021.
140. Yeung M, Sala E, Schönlieb C-B, et al. Focus U-Net: a novel dual attention-gated CNN for polyp segmentation during colonoscopy. Comput Biol Med 2021;137:104815.
141. Wang Y, Feng Z, Song L, et al. Multiclassification of endoscopic colonoscopy images based on deep transfer learning. Comput Math Methods Med 2021.
142. Yang C, Guo X, Zhu M, et al. Mutual-prototype adaptation for cross-domain polyp segmentation. IEEE J Biomed Health Inform 2021;25:3886-97.
143. Kim T, Lee H, Kim D. UACANet: uncertainty augmented context attention for polyp segmentation. In: Proceedings of the 29th ACM International Conference on Multimedia. Available at: https://dl.acm.org/doi/abs/10.1145/3474085.3475375. Accessed October 24, 2022.
144. Dong B, Wang W, Fan D-P, et al. Polyp-PVT: polyp segmentation with pyramid vision transformers. arXiv preprint arXiv:2108.06932 2021.
145. Sang DV, Chung TQ, Lan PN, et al. Ag-curesnest: a novel method for colon polyp segmentation. arXiv preprint arXiv:2105.00402 2021.
146. Patel K, Bur AM, Wang G. Enhanced u-net: a feature enhancement network for polyp segmentation. In: 2021 18th Conference on Robots and Vision. Proc Int Robot Vis Conf 2021;2021:181–8.
147. Wei J, Hu Y, Zhang R, et al. Shallow attention network for polyp segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021.
148. Lai H, Luo Y, Zhang G, et al. Toward accurate polyp segmentation with cascade boundary-guided attention. Vis Comput 2022:1-17.
149. Shin W, Lee MS, Han SW. COMMA: propagating complementary multi-level aggregation network for polyp segmentation. Appl Sci 2022;12:2114.
150. Yang K, Chang S, Tian Z, et al. Automatic polyp detection and segmentation using shuffle efficient channel attention network. Alexandria Eng J 2022;61:917-26.
151. Thanh NC, Long TQ. CRF-EfficientUNet: an improved UNet framework for polyp segmentation in colonoscopy images with combined asymmetric loss function and CRF-RNN layer. IEEE Access 2021;9:156987-7001.
152. Shen Y, Jia X, Meng MQ-H. HRENet: a hard region enhancement network for polyp segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021.
153. Ta N, Chen H, Lyu Y, et al. BLE-Net: boundary learning and enhancement network for polyp segmentation. Multimedia Syst 2022:1-14.
154. He F, Chen S, Li S, et al. Colonoscopic image synthesis for polyp detector enhancement via gan and adversarial training. In: 2021 IEEE 18th International Symposium on Biomedical Imaging. Available at: https://www.semanticscholar.org/paper/Colonoscopic-Image-Synthesis-For-Polyp-Detector-Via-He-Chen/febf3eafca47b1cd39d40e8d1362921fcbf2c3b8. Accessed October 24, 2022.
155. Dang T, Nguyen T, McCall J, et al. Ensemble learning based on classifier prediction confidence and comprehensive learning particle swarm optimisation for polyp localisation. arXiv preprint arXiv:2104.04832 2021.
156. Srivastava A, Chanda S, Jha D, et al. GMSRF-Net: an improved generalizability with global multi-scale residual fusion network for polyp segmentation. arXiv preprint arXiv:2111.10614 2021.
157. Safarov S, Whangbo TK, A-DenseUNet. adaptive densely connected UNet for polyp segmentation in colonoscopy images with atrous convolution. Sensors 2021;21:1441.
158. Qiu Z, Wang Z, Zhang M, et al. BDG-Net: boundary distribution guided network for accurate polyp segmentation. arXiv preprint arXiv:2201.00767 2022.
159. Shen Y, Jia X, Pan J, et al. APRNet: alternative prediction refinement network for polyp segmentation. Annu Int Conf IEEE Eng Med Biol Soc 2021;2021:3114-7.
160. Hung NB, Duc NT, Van Chien T, et al. AG-ResUNet++: an improved encoder-decoder based method for polyp segmentation in colonoscopy images. In: 2021 RIVF International Conference on Computing and Communication Technologies.
161. Wang M, An X, Li Y, et al. EMS-Net: enhanced multi-scale network for polyp segmentation. Annu Int Conf IEEE Eng Med Biol Soc 2021;2021:2936-9.
162. Huang D, Han K, Xi Y, et al. Multi-scale fusion attention network for polyp segmentation. In: International Conference on Neural Information Processing: Springer; 2021.
163. Thanh NC, Long TQ. Self-supervised visual feature learning for polyp segmentation in colonoscopy images using image reconstruction as pretext task. In: 2021 8th NAFOSTED Conference on Information and Computer Science. Available at: https://arxiv.org/pdf/2110.08776. Accessed October 24, 2022.
164. Chen Y, Liu Z, Shi Y. RP-Unet: a Unet-based network with RNNPool enables computation-efficient polyp segmentation. In: Sixth International Workshop on Pattern Recognition: International Society for Optics and Photonics; 2021.
165. Huang J, Zhou Z, Duan Y, et al. Information fusion for colorectal polyps medical image segmentation. Research Square, February 2022, p. 1-10.
166. Wang D, Hao M, Xia R, et al. MSB-Net: multi-scale boundary net for polyp segmentation. In: 2021 IEEE 10th Data Driven Control and Learning Systems Conference. Available at: https://ieeexplore.ieee.org/document/9455514. Accessed October 24, 2022.
167. Zizhao C. HarDNet-PSA: a high precision polyp detection neural network with attention module. In: 2021 IEEE International Conference on Emergency Science and Information Technology.
168. Yua C, Yana J, Lia X. Parallel Res2Net-based network with reverse attention for polyp segmentation. 2021.
169. Huang C-H, Wu H-Y, Lin Y-L. Hardnet-mseg: a simple encoder-decoder polyp segmentation neural network that achieves over 0.9 mean dice and 86 fps. arXiv preprint arXiv:2101.07172 2021.

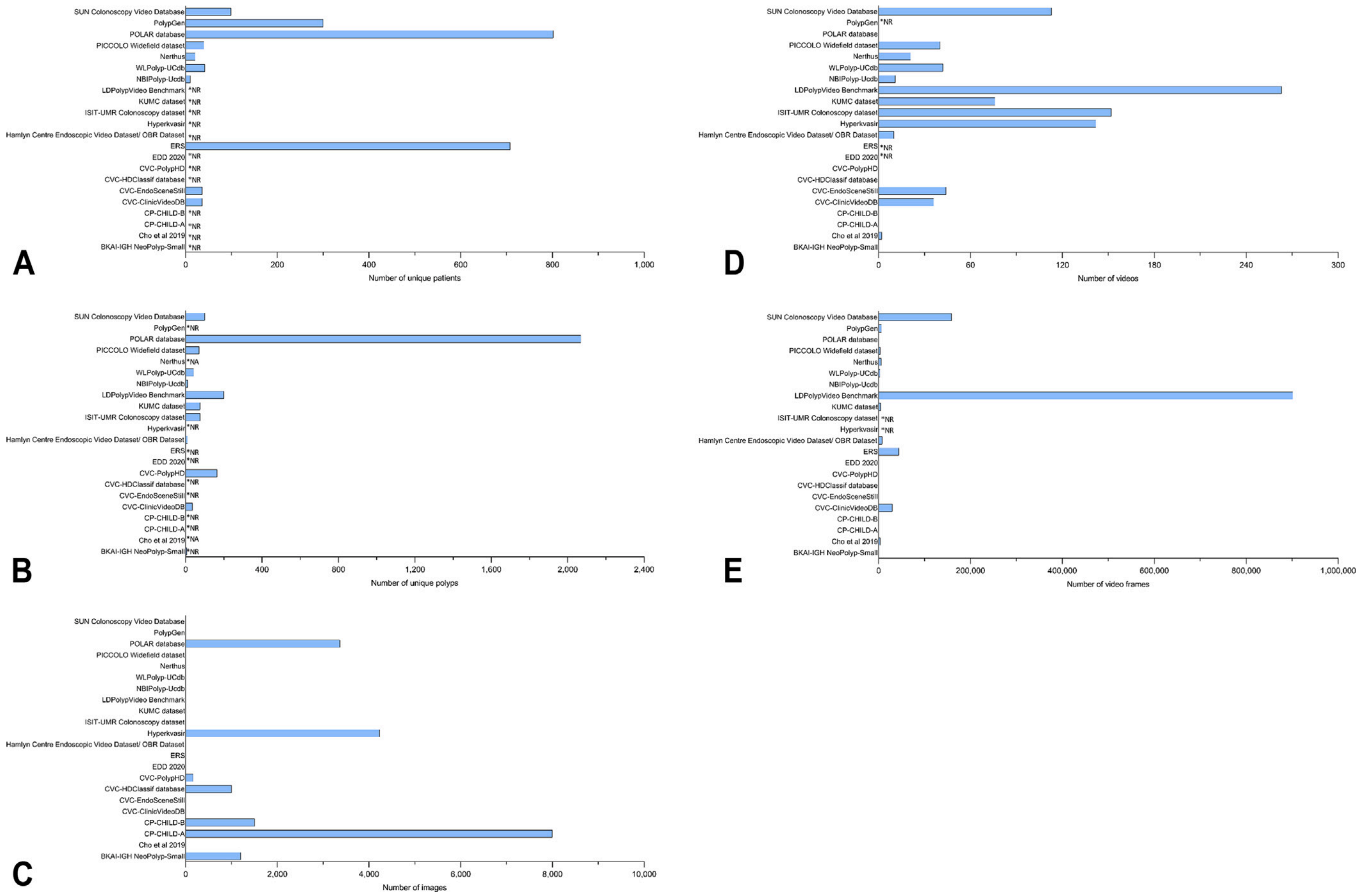
170. Wang S, Yin Y, Wang D, et al. An interpretable deep neural network for colorectal polyp diagnosis under colonoscopy. Knowl Based Syst 2021;234:107568.
171. Galdran A, Carneiro G, Ballester MAG. Double encoder-decoder networks for gastrointestinal polyp segmentation. In: Pattern recognition. ICPR International Workshops and Challenges. Cham, Switzerland: Springer International Publishing; 2021.
172. Sasmal P, Bhuyan M, Dutta S, et al. An unsupervised approach of colonic polyp segmentation using adaptive markov random fields. Pattern Recogn Lett 2022;154:7-15.
173. Xu J, Zhang Q, Yu Y, et al. Deep reconstruction-recoding network for unsupervised domain adaptation and multi-center generalization in colonoscopy polyp detection. Comput Methods Progr Biomed 2022;214:106576.
174. Erol T, Sarikaya D. An efficient polyp segmentation network. arXiv preprint arXiv:2203.04118 2022.
175. Nguyen T-C, Nguyen T-P, Diep G-H, et al. CCBANet: cascading context and balancing attention for polyp segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021.
176. Karmakar R, Nooshabadi S. Mobile-PolypNet: light-weight colon polyp segmentation network for low resources settings. 2022;8:1-9.
177. Ji G-P, Chou Y-C, Fan D-P, et al. Progressively normalized self-attention network for video polyp segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021.
178. Bardhi O, Sierra-Sosa D, Garcia-Zapirain B, et al. Deep learning models for colorectal polyps. Information 2021;12:245.
179. Hsu C-M, Hsu C-C, Hsu Z-M, et al. Colorectal polyp image detection and classification through grayscale images and deep learning. Sensors 2021;21:5995.
180. Nisha J, Gopi VP, Palanisamy P. Automated colorectal polyp detection based on image enhancement and dual-path CNN architecture. Biomed Signal Process Control 2022;73:103465.
181. Shen Z, Fu R, Lin C, et al. COTR: convolution in transformer network for end to end polyp detection. In: 2021 7th International Conference on Computer and Communications. Available at: https://arxiv.org/abs/2105.10925. Accessed October 24, 2022.
182. Eu CY, Tang TB, Lin C-H, et al. Automatic polyp segmentation in colonoscopy images using a modified deep convolutional encoder-decoder architecture. Sensors 2021;21:5630.
183. Sziová B, Nagy S, Kóczy LT. The effects of preprocessing on colorectal polyp detecting by fuzzy algorithm. In: Recent developments and the new direction in soft-computing foundations and applications: Springer; 2021. p. 347-57.
184. Nisha J, Gopi V, Palanisamy P. Automated polyp detection in colonoscopy videos using image enhancement and saliency detection algorithm. Biomed Eng Appl Basis Commun 2021.
185. Mahanty M, Bhattacharyya D, Midhunchakkaravarthy D. SRGAN assisted encoder-decoder deep neural network for colorectal polyp semantic segmentation. Rev Intell Artif 2021;35:395-401.
186. Gómez-Zuleta MA, Cano-Rosales DF, Bravo-Higuera DF, et al. Artificial intelligence techniques for the automatic detection of colorectal polyps. Rev Colomb Gastroenterol 2021;36:7-16.
187. Mahanty M, Bhattacharyya D, Midhunchakkaravarthy D. SU-NET based colorectal polyp segmentation from colon cancer morphology images. NVEO J 2021:649-59.
188. Wang D, Chen S, Chen Q, et al. AFP-Mask: anchor-free polyp instance segmentation in colonoscopy. IEEE J Biomed Health Inform 2022.
189. Zhang Z, Xiao L, Zhuang F, et al. An efficient polyp detection framework with suspicious targets assisted training. In: Chinese Conference on Pattern Recognition and Computer Vision: Springer; 2021.
190. Isidro AE, Fajardo AC, Hernandez AA. Exploration of various deep learning models for increased accuracy in automatic polyp detection. arXiv preprint arXiv:2203.04093 2022.
191. Ribeiro J, Nóbrega S, Cunha A. Polyps detection in colonoscopies. Proc Comput Sci 2022;196:477-84.
192. Kwon J, Choi K. Weakly supervised attention map training for histological localization of colonoscopy images. Annu Int Conf IEEE Eng Med Biol Soc 2021;2021:3725-8.
193. Mahato SJ, Banik D, Bhattacharjee D. Exploring hand-crafted features and transfer learning for polyp segmentation. In: International Conference on Computational Intelligence in Communications and Business Analytics: Springer; 2021.
194. Bae S-H, Yoon K-J. Polyp detection via imbalanced learning and discriminative feature learning. IEEE Trans Med Imag 2015;34:2379-93.
195. Figueiredo PN, Figueiredo IN, Pinto L, et al. Polyp detection with computer-aided diagnosis in white light colonoscopy: comparison of three different methods. Endosc Int Open 2019;7:E209-15.
196. Lee JY, Jeong J, Song EM, et al. Real-time detection of colon polyps during colonoscopy using deep learning: systematic validation with four independent datasets. Sci Rep 2020;10:1-9.
197. Podlasek J, Heesch M, Podlasek R, et al. Real-time deep learning-based colorectal polyp localization on clinical video footage achievable with a wide array of hardware configurations. Endosc Int Open 2021;9:E741-8.
198. Cai L, Beets-Tan R, Benson S. An improved automatic system for aiding the detection of colon polyps using deep learning. In: 2021 IEEE EMBS International Conference on Biomedical and Health Informatics. Available at: https://ieeexplore.ieee.org/document/9508557. Accessed October 24, 2022.
199. Sasmal P, Bhuyan M, Gupta S, et al. Detection of polyps in colonoscopic videos using saliency map-based modified particle filter. IEEE Trans Instrument Measure 2021;70:1-9.
200. Bernal J, Histace A. GTCreator: a flexible annotation tool for image-based datasets. 2019;14:191-201.
201. Sanchez-Montes C, Sanchez FJ, Bernal J, et al. Computer-aided prediction of polyp histology on white light colonoscopy using surface pattern analysis. Endoscopy 2019;51:261-5.
202. Ali S, Dmitrieva M, Ghatwary N, et al. Deep learning for detection and segmentation of artefact and disease instances in gastrointestinal endoscopy. Med Image Anal 2021.
203. Khadka R, Jha D, Hicks S, et al. Meta-learning with implicit gradients in a few-shot setting for medical image segmentation. Comput Biol Med 2022.
204. Cychnerski J, Dziubich T, Brzeski A. ERS: a novel comprehensive endoscopy image dataset for machine learning, compliant with the MST 3.0 specification. arXiv preprint arXiv:2201.08746, 2022.
205. Ye M, Giannarou S, Meining A, et al. Online tracking and retargeting with applications to optical biopsy in gastrointestinal endoscopic examinations. Med Image Anal 2016;30:144-57.
206. Nguyen DT, Lee MB, Pham TD, et al. Enhanced image-based endoscopic pathological site classification using an ensemble of deep learning models. Sensors 2020:20.
207. Ye M, Johns E, Walter B, et al. An image retrieval framework for real-time endoscopic image retargeting. Int J Comput Assist Radiol Surg 2017;12:1281-92.
208. Nagesh B, Kavya N. Polyp detection in endoscopy videos using cascaded classification. In: 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks. Available at: https://ieeexplore.ieee.org/abstract/document/9388443. Accessed October 24, 2022.
209. Armin MA, Barnes N, Khan S, et al. Unsupervised learning of endoscopy video frames' correspondences from global and local transformation. In: OR 2.0 context-aware operating theaters, computer assisted robotic endoscopy, clinical image-based procedures, and skin image analysis: Springer; 2018. p. 108-17.
210. Borgli H, Thambawita V, Smedsrud PH, et al. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. Sci Data 2020;7:283.
211. Pogorelov K, Randel KR, Griwodz C, et al. Kvasir: a multi-class image dataset for computer aided gastrointestinal disease detection. In: Proceedings of the 8th ACM on Multimedia Systems Conference. Available at: https://dl.acm.org/doi/10.1145/3083187.3083212. Accessed October 24, 2022.

212. Jha D, Smedsrud PH, Riegler MA, et al. Kvasir-seg: a segmented polyp dataset. In: International Conference on Multimedia Modeling: Springer; 2020. Available at: https://arxiv.org/abs/1911.07069. Accessed October 24, 2022.
213. Ahmed A, Ali M. Generative adversarial networks for automatic polyp segmentation. Available at: https://arxiv.org/abs/2012.06771. Accessed October 24, 2022.
214. Alam S, Tomar NK, Thakur A, et al. Automatic polyp segmentation using U-Net-ResNet50. Available at: https://arxiv.org/abs/2012.15247. Accessed October 24, 2022.
215. Ali SMF, Khan MT, Haider SU, et al. Depth-wise separable atrous convolution for polyps segmentation in gastro-intestinal tract. 2020. Available at: https://ceur-ws.org/Vol-2882/paper31.pdf. Accessed October 24, 2022.
216. Banik D, Bhattacharjee D. Deep Conditional adversarial learning for polyp segmentation. 2020. Available at: https://ceur-ws.org/Vol-2882/paper22.pdf. Accessed October 24, 2022.
217. Batchkala G, Ali S. Real-time polyp segmentation using U-Net with IoU loss. 2020. Available at: https://ceur-ws.org/Vol-2882/paper30.pdf. Accessed October 24, 2022.
218. Jha D, Hicks SA, Emanuelsen K, et al. Medico multimedia task at mediaeval 2020: automatic polyp segmentation. Available at: https://arxiv.org/abs/2012.15244. Accessed October 24, 2022.
219. Ji G-P, Fan D-P, Zhou T, et al. Automatic polyp segmentation via parallel reverse attention network. Available at: https://ceur-ws.org/Vol-2882/paper29.pdf. Accessed October 24, 2022.
220. Kang J, Gwak J. KD-ResUNet++: automatic polyp segmentation via self-knowledge distillation. 2020. Available at: https://ceur-ws.org/Vol-2882/paper74.pdf. Accessed October 24, 2022.
221. Khadka R. Transfer of knowledge: fine-tuning for polyp segmentation with attention. Medico 2020. Available at: https://ceur-ws.org/Vol-2882/paper15.pdf. Accessed October 24, 2022.
222. Nathan S, Ramamoorthy S. Efficient supervision net: polyp segmentation using EfficientNet and attention unit. 2020. Available at: https://www.slideshare.net/multimediaeval/efficient-supervision-net-polyp-segmentation-using-efficientnet-and-attention-unit. Accessed October 24, 2022.
223. Nguyen T-P, Nguyen T-C, Diep G-H, et al. HCMUS at medico automatic polyp segmentation task 2020: PraNet and ResUnet++ for polyps segmentation. 2020. Available at: https://ceur-ws.org/Vol-2882/paper47.pdf. Accessed October 24, 2022.
224. Obukhova NA, Motyko AA, Pozdeev AA. Two-stage method for polyps segmentation in endoscopic images. In: Computer Vision in Control Systems 6: Springer; 2020. p. 93-106.
225. Poudel S, Lee S-W. Automatic polyp segmentation using channel-spatial attention with deep supervision. 2020. Available at: https://ceur-ws.org/Vol-2882/paper24.pdf. Accessed October 24, 2022.
226. Puyal JG-B, Bhatia KK, Brandao P, et al. Endoscopic polyp segmentation using a hybrid 2D/3D CNN. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2020. Available at: https://www.researchgate.net/publication/345489385_Endoscopic_Polyp_Segmentation_Using_a_Hybrid_2D3D_CNN. Accessed October 24, 2022.
227. Trinh Q-H, Nguyen M-V, Huynh T-G, et al. Refined deep neural network and U-Net for polyps segmentation. arXiv preprint arXiv:2105.14848 2021. Available at: https://arxiv.org/abs/2105.14848. Accessed October 24, 2022.
228. Sasmal P, Paul A, Bhuyan M, et al. An automated framework for detection, localization, and classification of colonic polyp using deep learning. 2021. Available at: https://www.researchgate.net/publication/354751838_An_Automated_Framework_for_Detection_Localization_and_Classification_of_Colonic_Polyp_using_Deep_Learning. Accessed October 24, 2022.
229. Jha D, Ali S, Tomar NK, et al. Real-time polyp detection, localization and segmentation in colonoscopy using deep learning. IEEE Access 2021;9:40496-510.
230. Mohapatra S, Nayak J, Mishra M, et al. Wavelet transform and deep convolutional neural network-based smart healthcare system for gastrointestinal disease detection. Interdisc Sci Comput Life Sci 2021;13:212-28.
231. Chang Y-Y, Li P-C, Chang R-F, et al. Development and validation of a deep learning-based algorithm for colonoscopy quality assessment. Surg Endosc 2022:1-10.
232. Sutton RT, Zaïane OR, Goebel R, et al. Artificial intelligence enabled automated diagnosis and grading of ulcerative colitis endoscopy images. Sci Rep 2022;12:1-10.
233. Hicks S, Jha D, Thambawita V, et al. MedAI: transparency in medical image segmentation. Nord Mach Intell 2021;1:1-4.
234. Thambawita V, Strümke I, Hicks SA, et al. Impact of image resolution on deep learning performance in endoscopy image classification: an experimental study using a large dataset of endoscopic images. Diagnostics 2021;11:2183.
235. Tomar NK, Jha D, Ali S, et al. DDANet: dual decoder attention network for automatic polyp segmentation. In: International Conference on Pattern Recognition: Springer; 2021. Available at: https://arxiv.org/abs/2012.15245. Accessed October 24, 2022.
236. Tzavara NP, Singstad B-J. Transfer learning in polyp and endoscopic tool segmentation from colonoscopy images. Nord Mach Intell 2021;1:32-4.
237. Gjestang HL, Hicks SA, Thambawita V, et al. A self-learning teacher-student framework for gastrointestinal image classification. In: 2021 IEEE 34th International Symposium on Computer-Based Medical Systems. Available at: https://home.simula.no/∼paalh/students/HenrikGjestang-master.pdf. Accessed October 24, 2022.
238. Ayyoubi Nezhad S, Khatibi T, Sohrabi M. Proposing novel data analytics method for anatomical landmark identification from endoscopic video frames. J Healthc Eng 2022;2022:8151177.
239. Chou Y. Automatic polyp and instrument segmentation in MedAI-2021. Nord Mach Intell 2021;1:17-9.
240. Celik N, Ali S, Gupta S, et al. EndoUDA: a modality independent segmentation approach for endoscopy imaging. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021.
241. Somani A, Singh D, Prasad D, et al. T-MIS: transparency adaptation in medical image segmentation. Nord Mach Intell 2021;1:11-3.
242. Mohapatra S, Pati GK, Swarnkar T. Efficiency of transfer learning for abnormality detection using colonoscopy images: a critical analysis. In: 2022 IEEE Fourth International Conference on Advances in Electronics: Computers and Communications; 2022. Available at: https://ieeexplore.ieee.org/abstract/document/9716610. Accessed October 24, 2022.
243. Attallah O, Sharkas M. GASTRO-CADx: a three stages framework for diagnosing gastrointestinal diseases. PeerJ Comput Sci 2021;7:e423.
244. Öztürk Ş, Özkaya U. Residual LSTM layered CNN for classification of gastrointestinal tract diseases. J Biomed Inform 2021;113:103638.
245. Afify HM, Mohammed KK, Hassanien AE. An improved framework for polyp image segmentation based on SegNet architecture. Int J Imag Syst Technol 2021;31:1741-51.
246. Dulf E-H, Bledea M, Mocan T, et al. Automatic detection of colorectal polyps using transfer learning. Sensors 2021;21:5704.
247. Ahmed A. Classification of gastrointestinal images based on transfer learning and denoising convolutional neural networks. In: Proceedings of International Conference on Data Science and Applications: Springer; 2022.
248. Dalju HB, Rushdi MA, Morsy A. Superpixel-based segmentation and classification of gastrointestinal landmarks and diseases. In: 2021 IEEE Biomedical Circuits and Systems Conference. Available at: https://ieeexplore.ieee.org/abstract/document/9645002. Accessed October 24, 2022.
249. Hasnat R, Al Mamun A, Musha A, et al. Computer-aided polyp removal detection in endoscopic images. In: 2021 3rd International Conference on Electrical & Electronic Engineering. Available at: https://

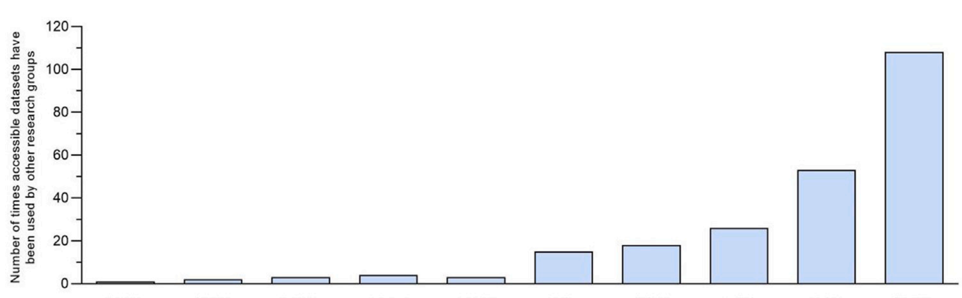
ieeexplore.ieee.org/abstract/document/9718791. Accessed October 24, 2022.
250. Dulf E-H, Muresan CI, Mocan T, et al. Computer-aided diagnosis system for colorectal cancer. In: 2021 25th International Conference on System Theory, Control and Computing.
251. Lumini A, Nanni L, Maguolo G. Deep ensembles based on stochastic activation selection for polyp segmentation. Available at: https://arxiv.org/abs/2104.00850. Accessed October 24, 2022.
252. Branch MV, Carvalho AS. Polyp segmentation in colonoscopy images using u-net-mobilenetv2. Available at: https://arxiv.org/abs/2103.15715. Accessed October 24, 2022.
253. Wu H, Zhong J, Wang W, et al. Precise yet efficient semantic calibration and refinement in ConvNets for real-time polyp segmentation from colonoscopy videos. In: Proceedings of the AAAI Conference on Artificial Intelligence 2021;35:2916-24.
254. Tomar NK. Automatic polyp segmentation using fully convolutional neural network. Available at: https://arxiv.org/abs/2101.04001. Accessed October 24, 2022.
255. Bhattacharya D, Betz C, Eggert D, et al. Self-supervised u-net for segmenting flat and sessile polyps. arXiv preprint arXiv:2110.08776 2021.
256. Sushma B, Raghavendra C, Prashanth J. CNN based U-Net with modified skip connections for colon polyp segmentation. In: 2021 5th International Conference on Computing Methodologies and Communication. Available at: https://ieeexplore.ieee.org/document/9418037. Accessed October 24, 2022.
257. Polat G, Işık Polat E, Kayabay K, et al. Polyp detection in colonoscopy images using deep learning and bootstrap aggregation. 2021. Available at: https://open.metu.edu.tr/handle/11511/89910. Accessed October 24, 2022.
258. Ali S, Tomar NK. Iterative deep learning for improved segmentation of endoscopic images. Nord Mach Intell 2021;1:38-40.
259. Ou S, Gao Y, Zhang Z, et al. Polyp-YOLOv5-Tiny: a lightweight model for real-time polyp detection. In: 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence.
260. Rahim T, Musaddiq A, Kim DS. Deep learning-based colorectal cancer detection in endoscopic images. 2021:522-523.
261. Thanh NC. Colonoscopy image classification using self-supervised visual feature learning. J Milit Sci Technol 2021:3-13.
262. Zhou X, Bai Z, Lu Q, et al. Colorectal polyp segmentation based on group convolution and transformer. In: 2021 China Automation Congress.
263. Wan J, Chen B, Yu Y. Polyp Detection from colorectum images by using attentive YOLOv5. Diagnostics 2021;11:2264.
264. Thambawita V, Hicks SA, Halvorsen P, et al. Divergentnets: medical image segmentation by network ensemble. arXiv preprint arXiv:2107.00283 2021. Available at: https://arxiv.org/abs/2107.00283. Accessed October 24, 2022.
265. Mesejo P, Pizarro D, Abergel A, et al. Computer-aided classification of gastrointestinal lesions in regular colonoscopy. IEEE Trans Med Imag 2016;35:2051-63.
266. Billah M, Waheed S. Gastrointestinal polyp detection in endoscopic images using an improved feature extraction method. Biomed Eng Lett 2018;8:69-75.
267. Billah M, Waheed S, Rahman MM. An automatic gastrointestinal polyp detection system in video endoscopy using fusion of color wavelet and convolutional neural network features. Int J Biomed Imag 2017;2017:9545920.
268. Cincar K, Sima I. Machine learning algorithms approach for gastrointestinal polyps classification. In: 2020 International Conference on INnovations in Intelligent SysTems and Applications. 2020. Available at: https://ieeexplore.ieee.org/abstract/document/9194659. Accessed October 24, 2022.
269. Mostafiz R, Hasan M, Hossain I, et al. An intelligent system for gastrointestinal polyp detection in endoscopic video using fusion of bidimensional empirical mode decomposition and convolutional neural network features. Int J Imag Syst Technol 2020;30:224-33.
270. Sasmal P, Bhuyan M, Sonowal S, et al. Improved endoscopic polyp classification using GAN generated synthetic data augmentation. In: 2020 IEEE Applied Signal Processing Conference. Available at: https://ieeexplore.ieee.org/document/9276732. Accessed October 24, 2022.
271. Sierra F, Gutiérrez Y, Martínez F. An online deep convolutional polyp lesion prediction over narrow band imaging (NBI). Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:2412-5.
272. Vispute M, Bhandari SH. Automated polyp classification of gastroenteric lesion in colonoscopy videos. In: 2018 5th International Conference on Signal Processing and Integrated Networks.
273. Wang Y, Chen X, Cai H, et al. Effect of various image information in polyp classification by deep learning with small dataset. In: 2019 IEEE International Conference on Signal Information and Data Processing.
274. Wang Y, Pomeroy M, Cao W, et al. Polyp classification by Weber's Law as texture descriptor for clinical colonoscopy. In: Medical Imaging 2019: Computer-Aided Diagnosis: International Society for Optics and Photonics; 2019. Available at: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10950/109502V/Polyp-classification-by-Webers-Law-as-texture-descriptor-for-clinical/10.1117/12.2512756.short. Accessed October 24, 2022.
275. Zhang R, Zheng Y, Mak TW, et al. Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain. IEEE J Biomed Health Inform 2017;21:41-7.
276. Bora K, Bhuyan M, Kasugai K, et al. Computational learning of features for automated colonic polyp classification. Sci Rep 2021;11:1-16.
277. Singh D, Singh B. Effective and efficient classification of gastrointestinal lesions: combining data preprocessing, feature weighting, and improved ant lion optimization. J Ambient Intell Human Comput 2021;12:8683-98.
278. Sasmal P, Bhuyan M, Iwahori Y, et al. Colonoscopic polyp classification using local shape and texture features. IEEE Access 2021;9:92629-39.
279. Rahman MM, Wadud MAH, Hasan MM. Computerized classification of gastrointestinal polyps using stacking ensemble of convolutional neural network. Inform Med Unlocked 2021;24:100603.
280. Eriyanti NA, Sigit R, Harsono T. Classification of colon polyp on endoscopic image using support vector machine. In: 2021 International Electronics Symposium.
281. Wang Q, Che H, Ding W, et al. Colorectal polyp classification from white-light colonoscopy images via domain alignment. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021. Available at: https://arxiv.org/abs/2108.02476. Accessed October 24, 2022.
282. Sierra-Jerez F, Martínez F. A deep representation to fully characterize hyperplastic, adenoma, and serrated polyps on narrow band imaging sequences. Health Technol 2022:1-13.
283. Sima I, Cincar K. Transfer learning-based classification of gastrointestinal polyps. In: 2021 IEEE 21st International Conference on Bioinformatics and Bioengineering. Available at: https://ieeexplore.ieee.org/document/9635497. Accessed October 24, 2022.
284. Doğan RS, Yılmaz B. Comparison of deep learning and conventional machine learning methods for classification of colon polyp types. EuroBiotech J 2021;5:34-42.
285. Ma Y, Chen X, Cheng K, et al. LDPolypVideo benchmark: a large-scale colonoscopy video dataset of diverse polyps. In: International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer; 2021. Available at: https://link.springer.com/chapter/10.1007/978-3-030-87240-3_37. Accessed October 24, 2022.
286. Figueiredo IN, Pinto L, Figueiredo PN, et al. Unsupervised segmentation of colonic polyps in narrow-band imaging data based on manifold representation of images and Wasserstein distance. Biomed Signal Process Control 2019;53.
287. Pogorelov K, Randel KR, de Lange T, et al. Nerthus: a bowel preparation quality video dataset. In: Proceedings of the 8th ACM on Multimedia Systems Conference. Available at: https://dl.acm.org/doi/10.1145/3083187.3083216. Accessed October 24, 2022.
288. Zhu Y, Xu Y, Chen W, et al. A CNN-based cleanliness evaluation for bowel preparation in colonoscopy. In: 2019 12th International Congress on Image and Signal Processing: BioMedical Engineering

and Informatics; 2019. Available at: https://ieeexplore.ieee.org/document/8965825. Accessed October 24, 2022.

289. Sánchez-Peralta LF, Pagador JB, Picón A, et al. PICCOLO white-light and narrow-band imaging colonoscopic dataset: a performance comparative of models and datasets. Appl Sci 2020;10:8501.
290. Pacal I, Karaman A, Karaboga D, et al. An efficient real-time colonic polyp detection with YOLO algorithms trained by using negative samples and large datasets. Comput Biol Med 2022;141:105031.
291. Houwen BBSL, Giotis I, Vleugels JLA,et al, on behalf of the POLAR study group. Development and validation of a computer-aided diagnosis system for characterization of diminutive colorectal polyps during live endoscopy: a multicentre study with benchmarking against screening endoscopists. 2022.
292. Ali S, Jha D, Ghatwary N, et al. PolypGen: a multi-center polyp detection and segmentation dataset for generalisability assessment. Available at: https://arxiv.org/abs/2106.04463. Accessed October 24, 2022.
293. Tomara NK, Ibtehazc N, Jhaa D, et al. Improving generalizability in polyp segmentation using ensemble convolutional neural network. 2021. Available at: https://ceur-ws.org/Vol-2886/paper5.pdf. Accessed October 24, 2022.
294. Galdran A, Carneiro G, Ballester. Multi-center polyp segmentation with double encoder-decoder networks. 2021. Available at: https://ceur-ws.org/Vol-2886/paper1.pdf. Accessed October 24, 2022.
295. Misawa M, Kudo SE, Mori Y, et al. Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video). Gastrointest Endosc 2021;93:960-7.e3.
296. Itoh H, Oda M, Jiang K, et al. Binary polyp-size classification based on deep-learned spatial information. Int J Comput Assist Radiol Surg 2021;16:1817-28.



Supplementary Figure 1. **A,** Number of unique patients per database. **B,** Number of unique polyps per databases. **C,** Images per database. **D,** Videos per database. **E,** Video frames per database. *NA*, Not applicable; *NR*, not reported.



Supplementary Figure 2. Use of publicly available databases by other artificial intelligence research groups.

SUPPLEMENTARY TABLE 1. Metadata reporting checklist of colonoscopic imaging databases for artificial intelligence research

No.	Item	Derived from CLAIM checklist
1.	Statement on participant consent.	New item
2.	Statement on ethical approval.	New item
3.	Data collection period.	New item
4.	Prospective or retrospective data acquisition.	CLAIM item 5
5.	Number of participating centers.	New item
6.	Number and qualifications of endoscopists collecting data.	New item
7.	Eligibility criteria (eg, polyps/patients).	CLAIM item 5
8.	Deidentification methods: Describing methods by which data have been deidentified and how protected health information has been removed according to U.S. (Health Insurance Portability and Accountability Act), European (General Data Protection Regulation), or relevant laws.	CLAIM item 12
9.	Definition of ground truth reference standard: Included detailed, specific definitions of the ground truth annotation and classes, sufficient to allow replications.	CLAIM item 14
10.	Source of ground truth annotations, qualifications and preparation of annotators. Only applicable to polyp detection, polyp localization, polyp segmentation, and/or quality of colonoscopy databases.	CLAIM item 16
11.	Measurement of inter- and intraobserver variability: Methods to mitigate variability and/or resolve discrepancies. Only applicable to polyp detection, polyp localization, polyp segmentation, and/or quality of colonoscopy databases.	CLAIM item 18
12.	Patient demographics (eg, age/gender).	CLAIM item 34
13.	Polyp demographics (eg, distribution polyp histology type, size, and morphology).	CLAIM item 34
14.	Number of included patients.	CLAIM item 34
15.	Number of unique cases (polyps /bowel quality categories).	CLAIM item 34
16.	Number of images and videos (frames).	CLAIM item 34
17.	Imaging modality.	New item
18.	Brand endoscopy system.	New item
19.	Scope type.	New item
20.	File format.	New item
21.	Resolution.	New item

CLAIM, Checklist for Artificial Intelligence in Medical Imaging.

SUPPLEMENTARY TABLE 2. Data extraction form

No.	Items
1.	Database (title).
2.	Original reference database (author, journal, and year).
3.	Other references that used this database (authors and year).
4.	Access type (open access/open access with barriers, regulated access).
5.	If applicable: Which barriers?
6.	Link to database.
7.	If applicable: E-mail request.
8.	If applicable: E-mail response.
9.	Country of origin.
10.	Database publication date.
11.	Data collection period.
12.	Statement on participant consent.
13.	Statement on ethical approval.
14.	Setting data acquisition (prospective/retrospective).
15.	Number of centers collecting data.
16.	Number of endoscopists collecting data.
17.	Age patients (mean $\pm$ standard deviation).
18.	Gender (M/F).
19.	Inclusion criteria.
20.	Exclusion criteria.
21.	Deidentification methods: Describing methods by which data have been deidentified and how protected health information has been removed according to U.S. (Health Insurance Portability and Accountability Act), European (General Data Protection Regulation), or relevant laws.
22.	Definition of ground truth reference standard: Included detailed, specific definitions of the ground truth annotation, sufficient to allow replications.
23.	Measurement of inter- and intraobserver variability: Methods to mitigate variability and/or resolve discrepancies. Only applicable to polyp detection, polyp localization, polyp segmentation, and/or quality of colonoscopy databases.
24.	Database of images (Y/N).
25.	If applicable: Number of images.
26.	If applicable: Number of included unique patients.
27.	If applicable: Number of included unique polyps.
28.	If applicable: Number of images regarding polyps.
29.	If applicable: Number of images regarding quality.
30.	Image file format (eg, TIF, jpg, png).
31.	Image resolution (min w $\times$ h – max w $\times$ h).
32.	Brand and type of scope.
33.	Imaging modality (eg, narrow-band imaging/high-definition white-light endoscopy).
34.	Additional unlabeled images? (Y/N).
35.	If applicable: Number of unlabeled images.

(*continued on the next page*)

SUPPLEMENTARY TABLE 2. Continued

No.	Items
36.	Database of videos (Y/N).
37.	Number of videos.
38.	Number of frames in video.
39.	If applicable: Number of included unique patients.
40.	If applicable: Number of included unique polyps.
41.	If applicable: Number of videos regarding polyps.
42.	If applicable: Number of videos regarding quality.
43.	Video file format (eg, .mov/.avi).
44.	Video resolution (min w $\times$ h – max w $\times$ h).
45.	Brand and type of scope.
46.	Imaging modality (eg, narrow-band imaging/high-definition white-light endoscopy).
47.	Additional unlabeled videos? (Y/N).
48.	If applicable: Number of unlabeled videos.
49.	Findings: Polyp detection (Y/N).
50.	Type of detection (eg, file name bounding box, binary mask).
51.	Ground truth (eg, expert).
52.	Number of polyps.
53.	Findings: Polyp localization (Y/N) .
54.	Type of localization (eg, bounding box, binary mask).
55.	Ground truth (eg, expert).
56.	Number of polyps.
57.	Findings: Polyp segmentation (Y/N).
58.	Type of segmentation (eg, binary mask).
59.	Ground truth (eg, expert).
60.	Number of polyps.
61.	Findings: Polyp characterization (Y/N).
62.	Number of labels/classes .
63.	Names of labels/classes.
64.	Ground truth (eg, histopathology).
65.	Endoscopist diagnosis (Y/N).
66.	Findings: Bowel preparation (Y/N).
67.	Type of bowel preparation scale (Boston Bowel Preparation Score, Ottawa, Aronchik).
68.	Number of labels/classes.
69.	Names of labels/classes.
70.	Ground truth (eg, expert).
71.	Ground truth number of assessor.
72.	Number of cases (each class).
73.	Findings: Cecal intubation landmarks (Y/N).
74.	Number of labels/classes.
75.	Names of labels/classes.
76.	Ground truth (eg, expert).
77.	Number of cases.
78.	Findings: Rectum retroflexion (Y/N).
79.	Ground truth (eg, expert).
	Number of cases.

SUPPLEMENTARY TABLE 3. Characteristics of open access with barriers and regulated access Database

Database	Country of origin	Publication year	Access type	Type of regulations	Data access details
ACP-ColonDB530[2]	Hong Kong	2020	Regulated	Formal agreement	E-mail request corresponding author
Gong 2020[3]	China	2020	Regulated	Formal agreement	E-mail request corresponding author
Krenzer 2022[4]	Germany	2022	Open access with barriers	Statement that data are available but no response to request	E-mail request corresponding author
Lee 2020[5]	Korea	2020	Regulated	Institutional or ethical approval	E-mail request corresponding author
Lee 2022[6]	USA	2022	Regulated	Formal agreement	E-mail request corresponding author
Low 2021[7]	Canada	2021	Regulated	Formal agreement	E-mail request corresponding author
Lui 2019[8]	Hongkong	2019	Regulated	Formal agreement	E-mail request corresponding author
NNUC database[9]	China	2016	Open access with barriers	Untraceable download link because of chinese website	http://math.nenu.edu.cn/nnucdb/
Polydeep database (IISGS)[10]	Spain	2021	Open access with barriers	E-mail response data will soon be publicly available	https://www.iisgaliciasur.es/home/biobank-iisgs
Renner 2018[11]	Germany	2018	Regulated	Formal agreement	E-mail request corresponding author
Repici 2020[12]	Italy	2020	Regulated	Formal agreement	E-mail request corresponding author
Saito 2021[13]	Japan	2021	Regulated	Formal agreement	E-mail request corresponding author
Sinonquel 2021[14]	Belgium	2021	Regulated	Formal agreement	E-mail request corresponding author
Urban 2018[15]	USA	2018	Regulated	Formal agreement	E-mail request corresponding author
Wang 2018[16]	China	2018	Regulated	Formal agreement	E-mail request corresponding author
Wang 2019[17]	China	2019	Regulated	Formal agreement	E-mail request corresponding author
Wang 2020[18]	China	2020	Regulated	Formal agreement	E-mail request corresponding author
Yamada 2019[19]	Japan	2019	Regulated	Formal agreement	E-mail request corresponding author

BP, Bowel preparation; CI, linked color imaging; *PC*, polyp classification; *PD*, polyp detection; *PL*, polyp localization; *PS*, polyp segmentation; *NR*, not reported.
*Training data.
†Test data.

SUPPLEMENTARY TABLE 3. Continued

	No. of images			No. of videos (frames)		
Data application	Total	Polyp annotation	Quality of colonoscopy annotation	Total	Polyp annotation	Quality of colonoscopy annotation
PD, PL, PC	0	0	0	NR (188,600),* NR (33,346)†	NR (77,153),* NR (13,973)†	0
Cecal landmarks	21,428*	0	21,428*	236 (17,564)	0	236 (17,564)
PD, PL	0	0	0	NR (346,165)	NR (346,165)	0 (0)
PD, PL, PC	420,* 1338†	420,* 1338†	0	181 (80,575),* 22 (351,122)†	181 (80,575),* 22 (351,122)†	0
BP	73,304*	0	73,304*	321 (NR)†	0	321 (NR)†
CI, BP	0	0	0	35 (13,522)	0	35 (13,522)
PC	8000,* 76†	8000,* 76†	0	0	0	0
PD, PL, PS	800	800	0	0	0	0
PD, PL, PC	0	0	0	330 (28,576)	330 (28,576)	0
	0	0	0	1,315 (NR)	1054 (NR)	0
PD, PL, PC	893,* 186†	893,* 186†	0	0	0	0
PD, PL	0	0	0	840 (NR)*	840 (NR)*	0
Anatomic segment classification	9995,* 5121†	0	9995,* 5121†	0	0	0
PD, PL, PC	0	0	0	329 (131,619)*	329 (131,619)*	0
PD, PL	9313*	4760*	0	9 (44,947),* 11 (NR)†	9 (44,947),* 11 (NR)†	0
PD, PL, PC	5545,* 27,113†	3634,* 5541†	0	192 (1,133,397)†	192 (1,133,397)†	0
PD, PL, BP	0	0	0	1058 (NR)	0	1058 (NR)
PD, PL, BP	0	0	0	962 (NR)	0	962 (NR)
PD, PL	4087,* 4840†	1943,* 752†	0	NR (134,874)*	NR (134,874)*	0

SUPPLEMENTARY TABLE 4. Database citations of open access datasets

Database	No. of database citations (impact)		
	Cited	Used for training and/or internal validation	Used for external validation/testing
BKAI-IGH NeoPolyp-Small[20]	3	0	0
Cho[21]	4	0	0
CP-CHILD-A[22]	20	1[23]	0
CP-CHILD-B[22]	20	0	1[23]
CVC-ClinicVideoDB[24,25]	27	16[26-41]	4[42-45]
CVC-EndoSceneStill[46-48]	772	165[2,4,24,27-37,39,40,42-45,49-193]	19[24,32,34,50,51,63,73,79,82,99,108,116,174,194-199]
CVC-HDClassif database[200,201]	0	0	0
CVC-PolypHD[46]	131	3[4,39,40]	0
EDD 2020[202]	30	3[4,130,203]	0
ERS[204]	0	0	0
Hamlyn Centre Endoscopic Video Dataset[205]	43	4[117,206-208]	1[209]
HyperKvasir[210-212]	474	107[4,37,58-60,64,81,84,88,95,96,105,126,128-130,134,136-153,155,156-169,171,174-176,186,191,213-263]	4[135,145,197,264]
ISIT-UMR Colonoscopy dataset[265]	107	28[55,66,87,90,94,106,120,133,199,266-284]	1[228]
KUMC dataset [94,131,132]	30	2[192,280]	0
LDPolypVideo Benchmark [285]	0	0	0
NBIPolyp-Ucdb[286]	2	0	0
WLPolyp-UCdb[195]	23	0	0
Nerthus[287]	53	1[95]	1[288]
PICCOLO Widefield dataset[289]	11	2[129,290]	0
POLAR database[291]	0	0	0
PolypGen [292]	7	6[128,168,257,264,293,294]	0
SUN Colonoscopy video database[295]	48	5[4,168,198,228,290]	1[296]