SUN Colonoscopy Video Database

Update

- 2022 10/08 : Added supplementary data: Accuracy of the latest approved model (July 2022).
- 2022 07/11 : Revised notations of polyp location in Table 2.
- 2020 10/07 : Released SUN Colonoscopy Video Database.

Abstract

SUN (**S**howa **U**niversity and **N**agoya **U**niversity) Colonoscopy Video Database is the colonoscopy-video database for the evaluation of an automated colorectal-polyp detection. The database comprises of still images of videos, which are collected at the Showa University Northern Yokohama Hospital. Mori Laboratory, Graduate School of Informatics, Nagoya University developed this database. Every frame in the database was annotated by the expert endoscopists at Showa University.

Summary of database

The SUN database includes **49,136 polyp frames** taken from **different 100 polyps**, which were fully annotated with bounding boxes. **Non-polyp scenes of 109,554 frames** are also included in this database. The characteristics of the database are summarized in Tables 1-3.

In polyp-exsiting frames, each polyp is annotated with a bounding box as shown in Figure 1. The file formats of images, and bounding boxes are jpeg and a text file, respectively. In the text file, each row represents a bounding box of a polyp, that is, "Fielename min_Xcoordinate,min_Ycoordinate,max_Xcorrdinate,max_Ycoordinate,class_id", where class_id of 0 and 1 represent polyp and non-polyp frames, respectively. Here are an examples:

    polyp1_00001.jpg 50,100,150,200,0
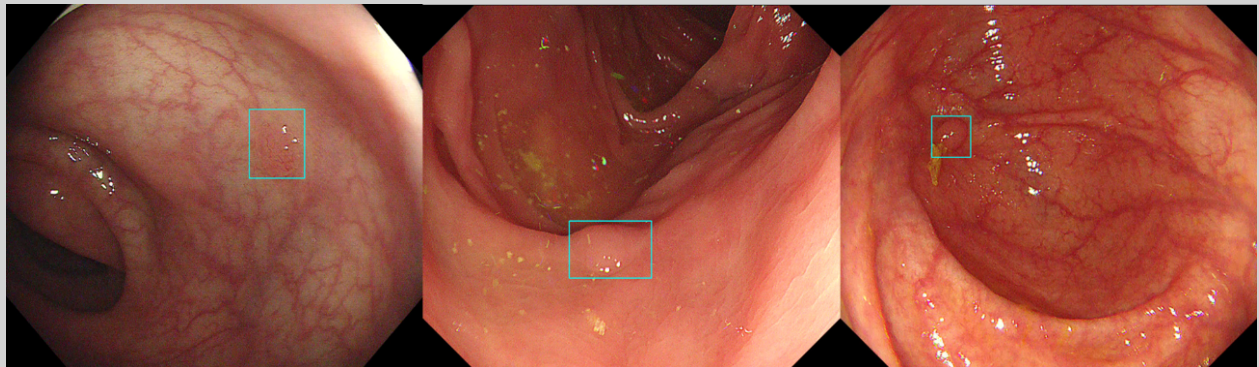   polyp1_00002.jpg 120,300,250,600,0
    ...

Figure 1: Examples of annotated images. The information of the bounding box is provided by a text file besides image files.

Table 1: Characteristics of the SUN database.

Patients registered as SUN database (n = 99)			
	Sex (Male / Female)		71 / 28
	Median Age (IQR)		69 (58 – 74)
Lesions registered as SUN database (n = 100)			
	Median Size (IQR) mm		5 (3 – 7)
	Number of diminutive polyp (≤5mm)		60
	Morphology (protruded / flat)		66 / 34
	Location (Right / Left / Rectum)		47 / 44 / 8
	Pathological diagnosis		
		Hyperplastic polyp	7
		Sessile serrated lesion	4
		Low grade adenoma	82
		Traditional serrated adenoma	2

| | | | | High grade adenoma | 4 |
| | | | | Invasive carcinoma | 1 |

Table 2: Breakdown of polyp samples of SUN database.

ID	Number of frames	Shape	Size	Location	Pathological diagnosis
1	527	Is	6mm	Cecum	Low-grade adenoma
2	1,313	Is	18mm	Rectum	High-grade adenoma
3	292	IIa	3mm	Ascending colon	Low-grade adenoma
4	80	Is	4mm	Sigmoid colon	Low-grade adenoma
5	930	IIa	3mm	Transverse colon	Low-grade adenoma
6	491	IIa	3mm	Sigmoid colon	Low-grade adenoma
7	315	IIa	6mm	Descending colon	Low-grade adenoma
8	256	Isp	12mm	Sigmoid colon	Low-grade adenoma
9	136	Is	4mm	Sigmoid colon	Low-grade adenoma
10	436	IIa	3mm	Transverse colon	Low-grade adenoma
11	113	IIa	5mm	Descending colon	Low-grade adenoma
12	538	Is	5mm	Rectum	Low-grade adenoma
13	479	Is	5mm	Transverse colon	Low-grade adenoma
14	1,183	IIa	3mm	Sigmoid colon	Low-grade adenoma
15	487	Is	5mm	Transverse colon	Low-grade adenoma
16	199	Is	4mm	Transverse colon	Low-grade adenoma
17	304	Is	4mm	Sigmoid colon	Low-grade adenoma
18	243	Is	2mm	Sigmoid colon	Hyperplastic polyp
19	96	IIa	3mm	Transverse colon	Low-grade adenoma
20	3159	IIa	3mm	Ascending colon	Low-grade adenoma
21	100	IIa	3mm	Sigmoid colon	Low-grade adenoma
22	314	IIa	2mm	Ascending colon	Low-grade adenoma
23	182	Ip	12mm	Ascending colon	Low-grade adenoma
24	973	Ip	15mm-	Sigmoid colon	Low-grade adenoma
25	338	Is	7mm	Sigmoid colon	Low-grade adenoma
26	370	Is	5mm	Descending colon	Low-grade adenoma
27	249	Is	5mm	Ascending colon	Hyperplastic polyp
28	195	Is	2mm	Transverse colon	Low-grade adenoma
29	377	Isp	13mm	Sigmoid colon	Low-grade adenoma
30	224	IIa	4mm	Sigmoid colon	Low-grade adenoma
31	183	Ip	12mm	Descending colon	Low-grade adenoma
32	981	Ip	15mm-	Ascending colon	Traditional serrated adenoma
33	594	Is	5mm	Sigmoid colon	Low-grade adenoma
34	245	Is	3mm	Ascending colon	Low-grade adenoma
35	1,212	Ip	15mm-	Sigmoid colon	High-grade adenoma
36	815	IIa	7mm	Sigmoid colon	Low-grade adenoma
37	448	Is	7mm	Transverse colon	Low-grade adenoma
38	509	Is	5mm	Ascending colon	Low-grade adenoma
39	713	IIa	13mm	Ascending colon	Low-grade adenoma
40	159	IIa	5mm	Transverse colon	Low-grade adenoma
41	108	IIa	3mm	Rectum	Low-grade adenoma
42	268	Is	7mm	Transverse colon	Low-grade adenoma
43	260	Isp	10mm	Ascending colon	Low-grade adenoma

44	745	IIa	5mm	Sigmoid colon	Low-grade adenoma
45	383	Is	3mm	Ascending colon	Low-grade adenoma
46	170	IIa	2mm	Transverse colon	Hyperplastic polyp
47	705	Is	5mm	Transverse colon	Low-grade adenoma
48	176	Is	3mm	Transverse colon	Low-grade adenoma
49	181	IIa	3mm	Transverse colon	Low-grade adenoma
50	740	Ip	10mm	Sigmoid colon	Low-grade adenoma
51	1,737	IIa(LST-NG)	15mm-	Cecum	Low-grade adenoma
52	207	IIa	6mm	Sigmoid colon	Low-grade adenoma
53	245	Is	4mm	Rectum	Hyperplastic polyp
54	345	Is	4mm	Sigmoid colon	Low-grade adenoma
55	700	Is	3mm	Ascending colon	Low-grade adenoma
56	248	Is	4mm	Sigmoid colon	Hyperplastic polyp
57	326	Is	5mm	Transverse colon	Low-grade adenoma
58	267	IIa	6mm	Transverse colon	Sessile serrated lesion
59	646	Isp	8mm	Sigmoid colon	Traditional serrated adenoma
60	146	IIa	8mm	Transverse colon	Low-grade adenoma
61	679	Isp	6mm	Ascending colon	Low-grade adenoma
62	351	Is	7mm	Ascending colon	Low-grade adenoma
63	632	Is	7mm	Rectum	Invasive cancer (T1b)
64	81	IIa	3mm	Sigmoid colon	Low-grade adenoma
65	222	IIa	3mm	Cecum	Low-grade adenoma
66	1,685	Is	6mm	Sigmoid colon	Low-grade adenoma
67	191	IIa	5mm	Transverse colon	Low-grade adenoma
68	1319	Is	15mm-	Rectum	High-grade adenoma
69	130	IIa	3mm	Descending colon	Low-grade adenoma
70	264	Ip	15mm-	Sigmoid colon	Low-grade adenoma
71	1,021	Is	4mm	Ascending colon	Low-grade adenoma
72	774	Is	5mm	Ascending colon	Low-grade adenoma
73	1,285	Is	3mm	Cecum	Low-grade adenoma
74	276	Isp	5mm	Sigmoid colon	Low-grade adenoma
75	343	Is	3mm	Transverse colon	Low-grade adenoma
76	343	Is	3mm	Cecum	Low-grade adenoma
77	215	Is	4mm	Ascending colon	Low-grade adenoma
78	267	Isp	12mm	Sigmoid colon	High-grade adenoma
79	76	Is	4mm	Descending colon	Low-grade adenoma
80	1,192	Is	10mm	Sigmoid colon	Low-grade adenoma
81	427	Is	6mm	Sigmoid colon	Low-grade adenoma
82	111	IIa	3mm	Sigmoid colon	Sessile serrated lesion
83	795	Isp	13mm	Rectum	Low-grade adenoma
84	218	Is	5mm	Descending colon	Low-grade adenoma
85	1,393	IIa	8mm	Ascending colon	Low-grade adenoma
86	257	IIa	4mm	Sigmoid colon	Low-grade adenoma
87	454	Is	3mm	Cecum	Low-grade adenoma
88	249	Is	4mm	Ascending colon	Low-grade adenoma
89	149	Ip	5mm	Descending colon	Low-grade adenoma
90	479	Is	10mm	Ascending colon	Sessile serrated lesion
91	1,061	IIa	13mm	Ascending colon	Low-grade adenoma

92	391	Is	7mm	Descending colon	Low-grade adenoma
93	452	Is	7mm	Descending colon	Low-grade adenoma
94	136	Is	6mm	Sigmoid colon	Low-grade adenoma
95	606	Isp	8mm	Sigmoid colon	Low-grade adenoma
96	301	Is	5mm	Sigmoid colon	Hyperplastic polyp
97	431	IIa	15mm-	Cecum	Sessile serrated lesion
98	170	IIa	4mm	Transverse colon	Low-grade adenoma
99	161	Is	5mm	Sigmoid colon	Low-grade adenoma
100	188	IIa	3mm	Rectum	Hyperplastic polyp

Table 3: Breakdown of non-polyp samples of SUN database.

ID	Number of frames	Lenth of each video (seconds)
1	9,961	332.0
2	10,073	335.8
3	7,152	238.4
4	14,635	487.8
5	7,916	263.9
6	17,046	511.4
7	5,636	169.1
8	2,568	85.6
9	9,522	317.4
10	7,086	236.2
11	4,832	161.1
12	6,799	226.6
13	6,328	210.9

[Back to top]

Supplementary data

Accuracy of the latest approved model (July 2022)

After the publication (Misawa M et al. Gastrointest Endosc 2021;93(4):960-967e3), we updated the EndoBRAIN-EYE (CADe). Therefore, we conducted re-analysis using the latest model to show current performance. The following tables show the latest performances for SUN Colonoscopy Video Database.

Table 4: The perfromances of the latest model and previously reported model.

	Latest model*		Reported model†	
	Percent (95% confidence interval)	n/N	Percent	n/N
Sensitivity (per-lesion)	98.0 (93.0-99.8)	98/100	98.0 (93.0-99.8)	98/100
Sensitivity (per-frame)	91.5 (91.2-91.7)	44,092/48212 ‡.§	90.5 (90.2-90.7)	44,472/49140 ‡
Specificity	98.2 (98.1-98.2)	90,068/91,764 ‡.§	93.7 (93.5-93.8)	88,075/94,039 ‡

*The latest model was regulatory approved in July 2022.
†Refer to the Misawa M et al. Gastrointest Endosc 2021;93(4):960-967e3.
‡The numbers of frames were inconsistent since the database was released with the removal of inappropriate frames after the paper publication.
§The frames which were identified as inappropriate frames by the CADe system were excluded from the analysis.

Table 5: Results of the positive videos (per-frame analysis).

Total number of frames	True positive*	False negative	False positive†	Unanalyzable frames‡
48,344	44,902	3,964	156	132

*Number of the frames in which the IoU between the predicted bounding box and ground truth more than or equals to 0.3.
†Number of the frames to which the trained model outputs bounding box but the IoU less than 0.3.
‡Number of the frames which were identified as inappropriate frame by the CADe system.

Table 6: Results of the negative videos (per-frame analysis)

Table 6: Results of the negative videos (per-frame analysis).

Total number of frames	False positive	True negative	Unanalyzable framese*
109,516	1,696	90,068	17,752

*Number of the frames which were identified as inappropriate frame by the CADe system.

Table 7: Results of the per-polyp analysis of the positive videos.

Number of polyp videos	True positive	False negative
100	98	2

[Back to top]

Terms of use

Copyright

All intellectual property rights, including copyrights, relating to the information contained in this database are held by Showa University Northern Yokohama Hospital and Mori Lab., Graduate School of Informatics, Nagoya University.

Intended use

This database is available for only non-commercial use in research or educational purpose. As long as you use the database for these purposes, you can edit or process images and annotations in this database. Without permission from Mori Lab., commercial use of this dataset is prohibited even after copying, editing, processing or any operations of this database. Please contact us for commercial use or if you are uncertain about the decision.

Citation

If you clearly indicate the following information, you are allowed to refer to or cite the part of this database. By downloading and using the SUN database you agree to cite this database in any publication based on research in which this database has been used.

- Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video). Masashi Misawa, Shin-ei Kudo, Yuichi Mori, Kinichi Hotta, Kazuo Ohtsuka, Takahisa Matsuda, Shoichi Saito, Toyoki Kudo, Toshiyuki Baba, Fumio Ishida, Hayato Itoh, Masahiro Oda, Kensaku Mori, Gastrointestinal Endoscopy, Vol. 93, Issue 4, pp. 960-967.e3, 2021. DOI: 10.1016/j.gie.2020.07.060
- SUN Colonoscopy Video Database. Hayato Itoh, Masashi Misawa, Yuichi Mori, Masahiro Oda, Shin-Ei Kudo, Kensaku Mori, 2020, http://amed8k.sundatabase.org/

Distribution

It is prohibited to sell, transfer, lend, lease, resell, distribute, etc., as it is, or copy, edit, or process this database, in whole or in part.

Personal information

The use of personal information sent to issue the download path of this database is limited to the notifications about this database and usage statistics of this database. Management of information follows the Nagoya University Privacy Policy. Note that this website uses Google Analytics with Cookie for analysis of its accesses. If you want to deny this analysis, please disable Cookie in your webbrowser.

Immunity

Showa University, Nagoya University, and Mori Lab. are not responsible for any damage caused by use and download of this database. We do not guarantee the accuracy, completeness, or usefulness of this database. If you violate any of the above terms, or at the discretion of Showa University, Mori Lab., or Nagoya University, you may not be allowed to use the database even after downloading. Showa University, Mori Lab, Nagoya University may arbitrarily revise this agreement without the approval of the registrant and this data user.

Contact

* If you agree to all the terms of use, please send a request e-mail to hitoh (a t) mori.m.is.nagoya-u.ac.jp.

[Back to top]