



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

OpenRC: An Open-Source Robotic Colonoscopy Framework for Multimodal Data Acquisition and Autonomy Research

Siddhartha Kapuria^{1*}, Mohammad Rafiee Javazm¹, Naruhiko Ikoma², Joga Ivatury³, Mohammad Ali Nasseri⁴, Nassir Navab⁴, and Farshid Alambeigi¹

¹ Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, 78712, TX, USA

skapuria@utexas.edu, farshid.alambeigi@austin.utexas.edu

² Department of Surgical Oncology, Division of Surgery, The University of Texas MD Anderson Cancer Center, Houston, 77030, TX, USA

³ Department of Surgery & Perioperative Care, Dell Medical School, The University of Texas at Austin, Austin, 78712, TX, USA

⁴ School of Medicine and Health, Technical University of Munich, Munich, Germany

Abstract. Colorectal cancer screening critically depends on colonoscopy, yet existing platforms offer limited support for systematically studying the coupled dynamics of operator control, instrument motion, and visual feedback. This gap restricts reproducible closed-loop research in robotic colonoscopy, medical imaging, and emerging vision-language-action (VLA) learning paradigms. To address this challenge, we present OpenRC, an open-source modular robotic colonoscopy framework that retrofits conventional scopes while preserving clinical workflow. The framework supports simultaneous recording of video, operator commands, actuation state, and distal tip pose. We experimentally validated motion consistency and quantified cross-modal latency across sensing streams. Using this platform, we collected a multimodal dataset comprising 1,894 teleoperated episodes (~19 hours) across 10 structured task variations of routine navigation, failure events, and recovery behaviors. By unifying open hardware and an aligned multimodal dataset, OpenRC provides a reproducible foundation for research in multimodal robotic colonoscopy and surgical autonomy. Both the hardware design and dataset are available at <https://github.com/artslab2019/openrc-robotic-colonoscopy>. The dataset has been contributed to the Open-H-Embodiment Initiative by NVIDIA.

Keywords: Robotic Colonoscopy · Multimodal Dataset · Surgical Robotics

1 Introduction

Colorectal cancer (CRC) is a major global health burden, ranking as the third most common cancer worldwide, with 1.9 million new cases and 935,000 deaths in 2020 [25,28]. Early detection is critical, as survival strongly depends on stage

* Corresponding author.

Table 1. Comparison of colonoscopy datasets and resources by available modalities

Dataset		Operator		Actuation State	6DoF Pose	Cross-Modal Alignment	Open Hardware
		Video	Actions				
Clinical datasets [4,12,6,5]	image/video	✓	✗	✗	✗	✗	✗
SimCol / SimCol3D [24,23]		✓	✗	✗	✓ (synthetic)	✓	✗
EndoSLAM [22]		✓	✗	✗	✓	✓	✗
EndoMapper [2]		✓	✗	✗	✓ (simulated subset)	✓	✗
C3VDv2 [10]		✓	✗	✗	✓	✓	Phantom STL/CAD
Copenhagen Colonoscopy Coordinate DB [9]		✗	✗	✗	✓	Single-modality	✗
OpenRC (Ours)		✓	✓	✓	✓	✓	✓

of cancer. Although colonoscopy is the gold standard and reduces mortality through removal of precancerous lesions [11], it still has adenoma miss rates of up to 34% [29]. These limitations arise from optical challenges (e.g., occlusion, blur), operator variability, and device ergonomics [15].

Investigating the literature, prior efforts in addressing colonoscopy challenges can broadly be categorized into two largely independent directions. The first emphasizes robotic and hardware innovation, including robotic platforms and motorized retrofits designed to improve dexterity, stability, and ergonomics (e.g., [3,7,16,17]). The second group focuses on AI-driven assistance through perception, classification, and SLAM, solely relying on colonoscopic video feed without considering instrument kinematics and control (e.g., [19,22,26]). Consequently, perception and control in these groups of efforts are typically studied in isolation.

Emerging autonomy-driven frameworks, particularly Vision-Language-Action (VLA) systems, require joint access to observations, actions, and system state. While such multimodal datasets are standard in broader robotics domains [21,27], comparable resources remain limited in colonoscopy. Existing open-source datasets predominantly provide video data and, in some cases, distal pose measurements, but rarely capture synchronized operator actions or internal actuation state (Table 1). Furthermore, there is currently no unified open-source hardware and software platform enabling reproducible closed-loop experimentation in colonoscopy. This lack of integrated platforms and datasets presents a fundamental barrier to reproducible research in robotic assistance and autonomy.

To address these challenges, we introduce *OpenRC*, an *open-source robotic colonoscopy research framework* for reproducible closed-loop experimentation and unified multimodal data collection. This low-cost platform retrofits conventional colonoscopes with modular actuation while preserving clinical workflow and synchronizes operator commands, actuation state, electromagnetic (EM)-tracked distal tip pose, and video. Unlike existing systems that are proprietary, expensive, or require specialized equipment (e.g. [16,17]), OpenRC provides an accessible, validated, and easy-to-adopt platform designed to interface with existing clinical hardware, rather than rebuilding the robotic platform

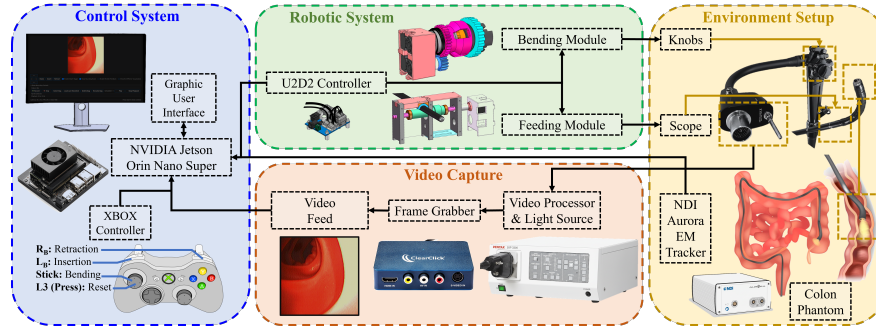


Fig. 1. Overview of the proposed OpenRC framework, including hardware interfaces, actuation, data acquisition, and sensing via EM tracking.

from scratch. Using this system, we conducted comprehensive hardware and data acquisition validation studies, and collected a multimodal dataset comprising 1,894 teleoperated episodes (~ 19 hours) spanning navigation and failure-recovery scenarios. The hardware design and dataset are made available at <https://github.com/artslab2019/openrc-robotic-colonoscopy> to support research in surgical autonomy, control-aware perception, and multimodal learning.

2 Methods

As shown in Fig. 1, to support research in computer vision, surgical autonomy, control-aware perception, and multimodal learning for colonoscopy procedures, the proposed open-source platform comprises four main modules: (1) a retrofitted robotic actuation unit for conventional colonoscopes; (2) integrated sensing and synchronized data logging; (3) configurable phantom testbeds; and (4) a ROS-based software architecture with a unified user interface. The entire framework (excluding the EM tracker) can be assembled for under \$5,000 USD, promoting accessibility and reproducibility. The following sections will describe these modules in detail.

2.1 Robotic Platform for Colonoscopy

The OpenRC system is designed to retrofit conventional commercially available colonoscopes, allowing for the development of a low-cost robotic system with repeated actuation and synchronized sensing without modifying the original device. In particular, it targets the three clinically relevant degrees-of-freedom (DoFs): insertion/retraction and distal-tip bending in the lateral and vertical directions. Without loss of generality, for this study, we use a PENTAX EC-3840LK colonoscope (Pentax Medical, Japan) as a representative model, although the system is designed to be compatible with other colonoscopes.

Bending Module. As shown in Fig. 2(a), the bending DoFs are motorized using two nested 3D-printed collets that transmit torque directly to the

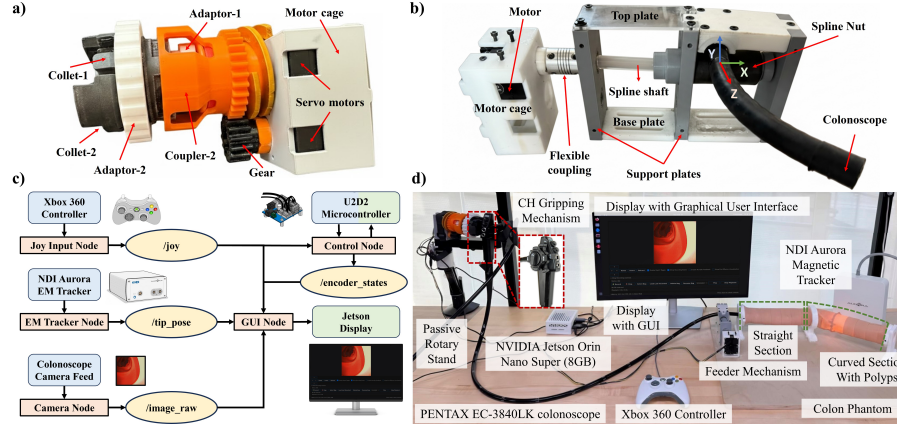


Fig. 2. OpenRC components: (a) bending module, (b) feeding module, (c) ROS 2 graph showing data streams, and (d) experimental setup for data collection.

existing control-handle (CH) knobs without modifying the colonoscope. Inspired by a drill collet, the concentric multi-jaw design enables secure gripping and rapid attachment/detachment. Two DYNAMIXEL XM540-W270-R servos (1:270 gearbox, 12-bit absolute encoder; 0.088° output resolution) provide actuation, delivering integrated sensing and sufficient holding torque, eliminating the need for the native lock. An additional 1:2 reduction in the outer collet improves command resolution. Components are fabricated from nylon-carbon fiber composite and ABS plastic to balance strength and manufacturability.

Feeding Module. As shown in Fig. 2(b), the feeding module uses a compact friction-based drive for insertion and retraction of the scope. A motor-driven urethane-sleeved feeder roller is used to provide traction, while a low-friction idler-ball guide applies an adjustable normal force via preload screws, ensuring consistent and controllable grip. The applied pressure can be tuned to accommodate different colonoscope diameters and to balance reliable advancement with safe force application. Any slippage at the roller interface remains observable and correctable in real time. The roller is actuated using a DYNAMIXEL XM430-W350-R servo (1:350 gearbox, 12-bit absolute encoder; 0.088° resolution). The feeding unit is mechanically independent of the bending module.

2.2 Software Architecture and Control Interface

As shown in Fig. 2(c), the system is built on a compact ROS 2 (Humble) stack running on an embedded NVIDIA Jetson Orin Nano Super (8 GB). Commands for bending and insertion actuation are issued to the DYNAMIXEL servos over an RS-485 daisy chain (ROBOTIS, South Korea) via a U2D2 interface and power hub board. Operator inputs are provided through an Xbox 360 controller and mapped to an action vector (`bend_x`, `bend_y`, `insertion`, `home`) normalized to

$[-1, 1]$, enabling simultaneous 3-DoF control and a reset-to-home command. The modular ROS 2 design enables straightforward integration of additional sensing modalities or autonomous control policies. Teleoperation and `rosvbag` logging execute in parallel to preserve control timing and ensure lossless recording.

2.3 Sensing and Logged Modalities

Through the ROS-based architecture described in Sec. 2.2, the OpenRC platform records multiple sensing modalities during colonoscope navigation via `rosvbag`, timestamped using a shared system clock. The following streams can be recorded:

Colonoscope Video. Video is captured at 30 frames per second at 383×396 pixels (native resolution) using the inbuilt camera at the distal tip of the colonoscope. As shown in Fig. 1, this feed is streamed via a Video to USB 1080P Capture device (ClearClick, USA) connected to a Video Processor and Light Source EPM 3300 (Pentax Medical, Japan) attached to the colonoscope.

Operator Commands. Operator intent is recorded as a four-dimensional action vector as described in Sec. 2.2, at a rate of 50 Hz. The normalized inputs provide a compact, abstracted representation of high-level operator commands independent of controller button mappings.

Robot State. In addition to operator commands, the joint-space actuation states are recorded at 50 Hz from the motor controllers. These signals comprise encoder-derived measurements for dual-axis distal bending and longitudinal translation, normalized to output shaft degrees-of-rotation to decouple encoder resolution from physical actuation. Importantly, due to the device’s inherent flexibility and environmental coupling, this joint-space representation does not uniquely map to the task-space distal tip pose.

Distal Tip Pose. Distal tip pose is measured at 6-DoF using an EM tracking system (NDI Aurora; Northern Digital Inc., Canada), represented as Cartesian position (m) and quaternion orientation. Since the colonoscope’s inherent flexibility and environmental interaction preclude a fixed spatial calibration, pose data are reported directly in the tracker’s coordinate frame.

2.4 Phantom Testbeds

We used two flexible colon phantoms: (i) a custom-built silicone phantom with embedded polyps and (ii) a commercially available training phantom. Importantly, the phantom-based setting enables acquisition of precise 6-DoF ground truth, which is not always feasible in clinical practice. Using both phantoms increases variation in geometry, surface appearance, and deformation characteristics without changing the robotic actuation and logging stack, improving the diversity of the resulting multimodal dataset.

Custom polyp phantom. Our custom phantom was designed to support both navigation and lesion-observation studies by permanently embedding realistic silicone polyps along the lumen. In brief, polyps were designed following the Paris classification [13,14], cast in silicone using 3D-printed molds, and integrated into a modular silicone colon segment during fabrication. The phantom

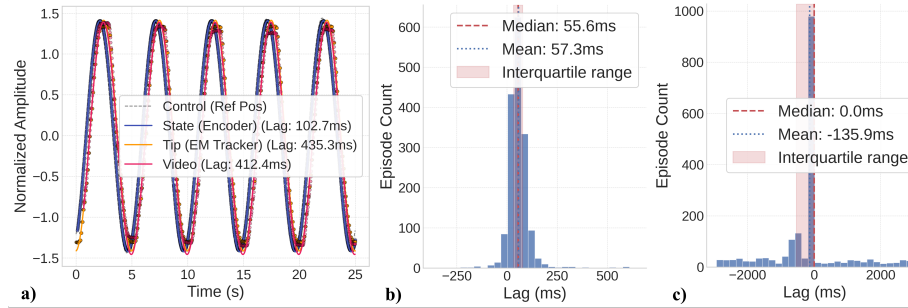


Fig. 3. Results for characterization and synchronization showing: (a) sinusoidal system response characterization, and histograms of estimated residual lag distributions for (b) Operator Action vs State, and (c) State vs Tip Position.

diameter was selected based on reported average human colon dimensions [1]. Multiple segments can be joined to adjust overall length and curvature.

Commercial flexible phantom. We also collected data in a commercially available flexible colonoscopy training phantom (Kyoto Kagaku, Japan).

3 System Characterization

As described in Sec. 2.3, OpenRC integrates operator input, joint-space motor states, EM-based distal tip tracking, and colonoscope video, which operate through independent acquisition pipelines. These readings are not inherently time-synchronized. Therefore, to assess hardware motion consistency and quantify cross-modal latency, we performed controlled excitation experiments. Specifically, we applied a low-frequency sinusoidal bending command (0.2 Hz) along the X axis in the feeding module frame (Fig. 2(b)). This excitation produced smooth, periodic motion observable across control commands, motor encoder measurements, distal tip position reported by the EM tracker, and motion estimates derived from optical flow (computed via the Lucas-Kanade method [18] and averaged across 5 manually selected representative keypoints in the video frame). For each modality, a sinusoidal model was fitted to the measured response, and relative phase differences were computed via least-squares regression. As shown in Fig. 3(a), the measured responses closely follow the commanded input, indicating consistent actuation. The estimated temporal offsets relative to the control action are approximately 102 ms for motor encoder state, 435 ms for EM tracking, and 412 ms for optical-flow-derived motion. These differences in offsets are consistent with the external data acquisition pipelines for each modality.

4 Multimodal Robotic Colonoscopy Dataset

Using the OpenRC framework, we collected a comprehensive multimodal dataset that captures data streams across the modalities described in Sec. 2.3. This

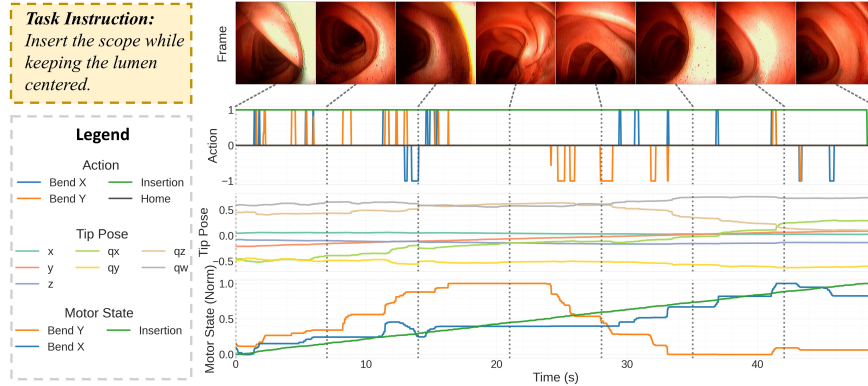


Fig. 4. Example episode with synchronized multimodal recordings from the robotic colonoscopy dataset. From top to bottom: colonoscopy video snapshots sampled from the same time axis as data streams; operator control actions; distal tip pose; robot state (normalized to $[-1, 1]$ per axis for ease of visualization).

section outlines the dataset structure, key features, technical validation, and potential applications in advancing robotic colonoscopy and autonomy research. To ensure accessibility, the dataset is saved in the LeRobot 2.1 format [8].

Episode Definition. The dataset is organized into episodes, where each episode corresponds to a single continuous teleoperated navigation trial performed within a colon phantom. Episodes include active manipulation, pauses, and adjustments, with durations varying by operator behavior and phantom geometry. For each episode, a short natural language instruction describing the intended task or navigation objective is recorded as metadata. Fig. 4 illustrates a representative teleoperated episode from the dataset for the insertion task, showing the temporal evolution across modalities.

Dataset Characteristics. The dataset comprises 1,894 episodes spanning ~ 19 hours of teleoperated navigation across 10 task variations focused on insertion–retraction and structured wall-scanning behaviors. Specifically, tasks 0–4 correspond to **insertion** with structured scanning objectives: **bottom wall** (0), **left wall** (1), **right wall** (2), **top wall** (3), and **lumen-centered navigation** (4). Tasks 5–9 mirror these objectives during **retraction**. As shown in Fig. 5(c), lumen-centered insertion and retraction tasks (tasks 4 and 9) predominate, reflecting their central role in scope manipulation, while wall-scanning tasks (tasks 0–3 and 5–8) are less frequent due to increased steering complexity. The dataset also includes 142 **failure** and 141 **recovery** episodes capturing common challenges such as lumen loss, wall contact, and fold engagement, along with corrective maneuvers. This enables structured study of error recovery in colonoscopic navigation. Episode-level statistics in Fig. 5(a–b) show a clear separation between routine navigation and failure/recovery segments. Navigation episodes

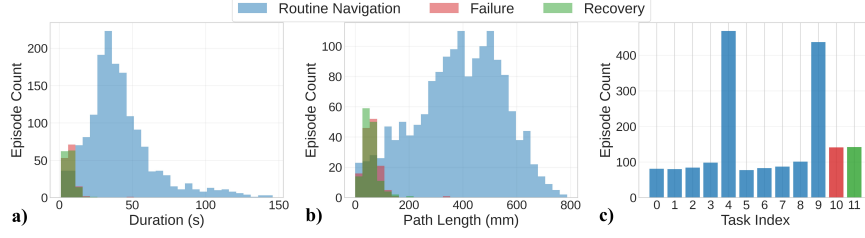


Fig. 5. Episode-level characteristics of the OpenRC Dataset showing distributions of (a) episode duration, (b) trajectory length, and (c) recorded task. Detailed task descriptions will be provided in the dataset repository.

are substantially longer and involve greater trajectory lengths, whereas failure and recovery segments are shorter and more localized in motion.

Data Synchronization. To ensure accurate cross-modal analysis and support typical VLA learning pipelines, all data streams are resampled to 30 Hz and calibrated using the offsets from Sec. 3, with the video feed as reference. To validate post-alignment stability, we estimated the residual temporal lag τ^* between modality pairs (x, y) , namely, operator action vs. actuation state, and actuation state vs. EM tip pose. To reduce sensitivity to absolute offsets, we extracted normalized, velocity-like representations. For a given signal x (and identically for y), we computed the finite difference Euclidean norm $v_t^x = \|x_t - x_{t-1}\|_2$, and applied min-max normalization as $\hat{v}^x = (v^x - \min(v^x)) / (\max(v^x) - \min(v^x))$. The residual offset was defined as the lag that maximizes the Pearson correlation coefficient ρ over a bounded symmetric window $[-\tau_{\max}, \tau_{\max}]$, formulated as $\tau^* = \arg \max_{\tau} \rho(\hat{v}_t^x, \hat{v}_{t+\tau}^y)$, where ρ was computed over the valid temporal overlap of the two signals. As shown in Fig. 3(b-c), residual lag distributions show temporal stability across modalities. Action vs. state (Fig. 3(b)) shows a tightly clustered median offset of 55.6 ms (~ 1.6 frames), well within acceptable limits for downstream policy learning. State vs. distal tip positions (Fig. 3(c)) center at 0.0 ms, validating the baseline controller-to-tracker alignment. The wider variance and fat tails in this case can be attributed to the continuum dynamics of the deformable environment, which can decouple distal motion from actuation.

5 Discussion and Conclusion

In this work, we introduced OpenRC, a low-cost ($< \$5,000$ USD) open-source robotic colonoscopy framework designed to retrofit conventional colonoscopes for reproducible closed-loop experimentation and unified multimodal data collection. The platform provides modular actuation of the clinically relevant DoFs, integrated multimodal sensing, and a ROS 2-based teleoperation and logging framework. Furthermore, we release a temporally synchronized multimodal dataset (in LeRobot 2.1 format) collected using this platform, comprising 1,894 episodes (~ 19 hours) of teleoperated procedures, including insertion, retraction, struc-

tured wall-scanning, and failure/recovery scenarios. Experimental validation confirmed reliable system actuation and yielded a stable post-alignment median residual lag of 55.6 ms between operator actions and actuation states, and 0.0 ms between motor states and EM-tracked distal tip poses. The collected synchronized action-state-pose-video trajectories provide a reproducible foundation for advancing research in surgical autonomy, control-aware perception, and multimodal learning in colonoscopy procedures. This dataset has been contributed to the Open-H-Embodiment Initiative by NVIDIA [20].

We acknowledge that the scope of our platform and dataset is currently limited by the controlled phantom environment used for data collection. Specifically, these *in vitro* setups do not comprehensively capture patient-specific anatomical variability and dynamic physiological motion. Despite this limitation, since the platform is modular and the dataset follows a standardized format, other researchers can collect compatible multimodal data using their own colonoscopes and phantoms, enabling community-driven expansion of this resource. In the future, we plan to test the system on more realistic colon phantoms and animal models. We will also leverage this platform and dataset to train and evaluate robust VLA models for autonomous colonoscopic navigation and error recovery.

Acknowledgments. This work was supported by the National Cancer Institute of the National Institutes of Health under Award Number R21CA280747, the IC² Institute, and Texas Robotics Seed Collaborative Funding at the University of Texas at Austin.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Alazmani, A., Hood, A., Jayne, D.G., Neville, A., Culmer, P.: Quantitative assessment of colorectal morphology: Implications for robotic colonoscopy. *Medical engineering & physics* **38** *2*, 148–54 (2016)
2. Azagra, P., Sostres, C., Ferrandez, Á., Riazuelo, L., Tomasini, C., Barbed, O.L., Morlana, J., Recasens, D., Batlle, V.M., Gómez-Rodríguez, J.J., Elvira, R., López, J., Oriol, C., Civera, J., Tardós, J.D., Murillo, A.C., Lanas, Á., Montiel, J.M.M.: Endomapper dataset of complete calibrated endoscopy procedures. *Scientific Data* **10** (2022)
3. Basha, S., Khorasani, M., Abdurahiman, N., Padhan, J., Baez, V.M., Al-Ansari, A.A., Tsiamyrtzis, P., Becker, A.T., Navkar, N.V.: A generic scope actuation system for flexible endoscopes. *Surgical Endoscopy* **38**, 1096–1105 (2023)
4. Bernal, J., Sánchez, F.J., Fernández-Esparrach, G., Gil, D., Rodríguez, C., Vilar-iño, F.: Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics* **43**, 99–111 (2015)
5. Biffi, C., Antonelli, G., Bernhofer, S., Hassan, C., Hirata, D., Iwatate, M., Maieron, A., Salvagnini, P., Cherubini, A.: REAL-Colon: a dataset for developing real-world ai applications in colonoscopy. *Scientific Data* **11** (2024)

6. Borgli, H., Thambawita, V., Smedsrud, P.H., Hicks, S., Jha, D., Eskeland, S.L., Randel, K.R., Pogorelov, K., Lux, M., Nguyen, D.T.D., Johansen, D., Griwodz, C., Stensland, H.K., Garcia-Ceja, E., Schmidt, P.T., Hammer, H.L., Riegler, M.A., Halvorsen, P., de Lange, T.: HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Scientific Data* **7**(1), 283 (2020)
7. Boškoski, I., Orlandini, B., Papparella, L.G., Matteo, M.V., De Siena, M., Pontecorvi, V., Costamagna, G.: Robotics and artificial intelligence in gastrointestinal endoscopy: updated review of the literature and state of the art. *Current Robotics Reports* **2**(1), 43–54 (2021)
8. Cadene, R., Alibert, S., Soare, A., Gallouedec, Q., Zouitine, A., Palma, S., Kooijmans, P., Aractingi, M., Shukor, M., Aubakirova, D., Russi, M., Capuano, F., Pascal, C., Choghari, J., Moss, J., Wolf, T.: LeRobot: State-of-the-art machine learning for real-world robotics in pytorch. <https://github.com/huggingface/lerobot> (2024)
9. Cold, K.M., Vamadevan, A., Heen, A., Vilmann, A.S., Bulut, M., Kovačević, B., Rasmussen, M., Konge, L., Svendsen, M.B.S.: Mapping the colon through the colonoscope’s coordinates – the copenhagen colonoscopy coordinate database. *Scientific Data* **12** (2025)
10. Golhar, M.V., Fretes, L.S.G., Ayers, L., Akshintala, V.S., Bobrow, T.L., Durr, N.J.: C3VDv2-Colonoscopy 3D video dataset with enhanced realism. *arXiv preprint arXiv:2506.24074* (2025)
11. Jemal, A., Thomas, A., Murray, T., Thun, M.J.: Cancer statistics, 2002. *CA: A Cancer Journal for Clinicians* **52** (2002)
12. Jha, D., Smedsrud, P.H., Riegler, M.A., Halvorsen, P., de Lange, T., Johansen, D., Johansen, H.D.: Kvasir-seg: A segmented polyp dataset. In: *MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020, Proceedings, Part II* 26. pp. 451–462 (2020)
13. Kaltenbach, T., Anderson, J.C., Burke, C.A., Dominitz, J.A., Gupta, S., Lieberman, D.A., Robertson, D.J., Shaukat, A., Syngal, S., Rex, D.K.: Endoscopic removal of colorectal lesions-recommendations by the us multi-society task force on colorectal cancer. *Gastroenterology* (2020)
14. Kara, O.C., Venkatayogi, N., Ikoma, N., Alambeigi, F.: A reliable and sensitive framework for simultaneous type and stage detection of colorectal cancer polyps. *Annals of Biomedical Engineering* **51**, 1499–1512 (2023)
15. Ko, C.W., Dominitz, J.A.: Complications of colonoscopy: magnitude and management. *Gastrointestinal endoscopy clinics of North America* **20** **4**, 659–71 (2010)
16. Kume, K., Sakai, N., Goto, T.: Development of a novel endoscopic manipulation system: the endoscopic operation robot ver.3. *Endoscopy* **47**, 815–819 (2015)
17. Lee, D.H., Cheon, B., Kim, J., Kwon, D.S.: EasyEndo robotic endoscopy system: Development and usability test in a randomized controlled trial with novices and physicians. *The International Journal of Medical Robotics and Computer Assisted Surgery* **17**(3), e2168 (2020)
18. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: *Proceedings of the 7th international joint conference on Artificial intelligence-Volume 2*. pp. 674–679. Morgan Kaufmann Publishers Inc. (1981)
19. Nie, Z., Xu, M., Wang, Z., Lu, X., Song, W.: A review of application of deep learning in endoscopic image processing. *Journal of Imaging* **10**(11), 275 (2024)
20. Open-H-Embodiment Consortium: Open-H-Embodiment: A large-scale dataset for enabling foundation models in medical robotics. <https://arxiv.org/abs/2604.21017> (2026)

21. Open X-Embodiment Collaboration: Open X-Embodiment: Robotic learning datasets and RT-X models. <https://arxiv.org/abs/2310.08864> (2023)
22. Ozyoruk, K.B., Gokceler, G.I., Bobrow, T.L., Coskun, G., Incetan, K., Almalioglu, Y., Mahmood, F., Curto, E., Perdigoto, L., Oliveira, M., et al.: EndoSLAM dataset and an unsupervised monocular visual odometry and depth estimation approach for endoscopic videos. *Medical image analysis* **71**, 102058 (2021)
23. Rau, A., Bano, S., Jin, Y., Azagra, P., Morlana, J., Kader, R., Sanderson, E., Matuszewski, B.J., Lee, J.Y., Lee, D.J., et al.: SimCol3D—3D reconstruction during colonoscopy challenge. *Medical Image Analysis* **96**, 103195 (2024)
24. Rau, A., Bhattarai, B., Agapito, L., Stoyanov, D.: Bimodal camera pose prediction for endoscopy. *IEEE Transactions on Medical Robotics and Bionics* **5**(4), 978–989 (2023)
25. Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F.: Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **71**, 209 – 249 (2021)
26. Tham, S., Koh, F.H., Ladlad, J., Chue, K.M., Centre, S.E., Lin, C.L., Teo, E.K., Foo, F.J.: The imitation game: a review of the use of artificial intelligence in colonoscopy, and endoscopists’ perceptions thereof. *Annals of Coloproctology* **39**(5), 385 (2023)
27. Walke, H., Black, K., Lee, A., Kim, M.J., Du, M., Zheng, C., Zhao, T., Hansen-Estruch, P., Vuong, Q., He, A., Myers, V., Fang, K., Finn, C., Levine, S.: Bridge-Data V2: A dataset for robot learning at scale. In: *Conference on Robot Learning (CoRL)* (2023)
28. Xi, Y., Xu, P.: Global colorectal cancer burden in 2020 and projections to 2040. *Translational Oncology* **14**(10), 101174 (2021)
29. Zhao, S.B., Wang, S., Pan, P., Xia, T., Chang, X., Yang, X., Guo, L., Meng, Q.Q., Yang, F., Qian, W., Xu, Z., Wang, Y., Wang, Z., Gu, L., Wang, R., Jia, F., Yao, J., Li, Z., Bai, Y.: Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: A systematic review and meta-analysis. *Gastroenterology* **156** 6, 1661–1674.e11 (2019)