

MUT#1 ♦ Bowel Cleanliness Grading

Task definition: Assess bowel preparation quality for colonoscopy using the Boston-Bowel-Preparation-Scale (BBPS) criteria.

Task type: Procedural quality control

Evaluation metrics: Accuracy

Category: CLS#66 ~ CLS#71

Dataset origin: DATA#11, DATA#32

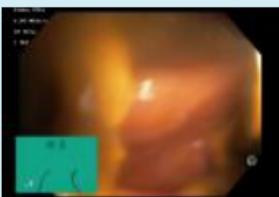
Number of images: Training 4,390 | Validation 733 | Testing 2,196

Number of VQA entries: Training 4,390 | Validation 733 | Testing 2,196

Task instruction templates:

- 1.<IMAGE>\nAssess the bowel preparation quality shown in the image: <OPTIONS>
- 2.<IMAGE>\nWhat category best describes the bowel preparation quality in the image? <OPTIONS>
- 3.<IMAGE>\nRate the quality of the bowel preparation in the image: <OPTIONS>
- 4.<IMAGE>\nDetermine the bowel preparation quality from the image: <OPTIONS>
- 5.<IMAGE>\nClassify the quality of bowel preparation depicted in the image: <OPTIONS>

Conversation example:



- ▶ **Human:** Classify the quality of bowel preparation depicted in the image: <A> Boston bowel preparation scale 1, Boston bowel preparation scale 2, <C> Boston bowel preparation scale 2-3, <D> Boston bowel preparation scale 0, <E>Boston bowel preparation scale 3, <F> Boston bowel preparation scale 0-1
- ▶ **Machine:** <D> Boston bowel preparation scale 0

MUT#2 ♦ Colonoscopy Completion Landmark Identification

Task definition: Identify anatomical landmarks for complete colon exam and navigation.

Task type: Procedural quality control

Evaluation metrics: Accuracy

Category: CLS#51, CLS#51, CLS#76

Dataset origin: DATA#10, DATA#12 ~ DATA#14

Number of images: Training 3,306 | Validation 551 | Testing 1,654

Number of VQA entries: Training 6,612 | Validation 1,102 | Testing 3,308

Task instruction templates:

- 1.<IMAGE>\nCan you identify whether the image includes anatomical landmarks, such as the cecum or ileocecal valve, that confirm the completion of a colonoscopy?
- 2.<IMAGE>\nAnalyze the image and determine if it shows key anatomical landmarks like the cecum or ileocecal valve, indicating the colonoscopy is complete.
- 3.<IMAGE>\nDoes the provided image depict the cecum or ileocecal valve, which are necessary landmarks for verifying a complete colonoscopy?
- 4.<IMAGE>\nEvaluate the image to confirm whether it displays critical anatomical landmarks, such as the cecum or ileocecal valve, for completing a colonoscopy.
- 5.<IMAGE>\nBased on the image, decide if the anatomical landmarks confirming colonoscopy completion (e.g., cecum or ileocecal valve) are visible.

Conversation example:



- **Human:** Evaluate the image to confirm whether it displays critical anatomical landmarks, such as the cecum or ileocecal valve, for completing a colonoscopy.
- **Machine:** Yes. This image shows cecum.

MUT#3 ♦ Rectum Retroflexion Identification

Task definition: Identify abnormality types in colonoscopy images from two to ten predefined options for diagnosis and treatment.

Task type: Procedural quality control

Evaluation metrics: Accuracy

Category: CLS#52

Dataset origin: DATA#11 ~ DATA#14

Number of images: Training 417 | Validation 70 | Testing 208

Number of VQA entries: Training 834 | Validation 140 | Testing 416

Task instruction templates:

- 1.<IMAGE>\nDoes this image show a retroflexed view of the rectum? Answer Yes or No.
- 2.<IMAGE>\nCan you confirm retroflexion in the rectum based on this image? Answer Yes or No.
- 3.<IMAGE>\nIs the observed perspective indicative of a retroflexed rectum view? Answer Yes or No.
- 4.<IMAGE>\nDoes the image confirm a retroflex examination? Answer Yes or No.
- 5.<IMAGE>\nCan retroflexion be identified in this rectal image? Answer Yes or No.

Conversation example:



- ▶ **Human:** Is the observed perspective indicative of a retroflexed rectum view? Answer Yes or No.
- ▶ **Machine:** Yes

MUT#4 ♦ Intervention Stage Recognition

Task definition: Determine the intervention state shown in colonoscopy images to assist clinical processes.

Task type: Procedural quality control

Evaluation metrics: Accuracy

Category: CLS#48 ~ CLS#51

Dataset origin: DATA#10 ~ DATA#13, DATA#15

Number of images: Training 1,983 | Validation 329 | Testing 991

Number of VQA entries: Training 1,983 | Validation 329 | Testing 991

Task instruction templates:

- 1.<IMAGE>\nWhat intervention phase does this image represent? <OPTIONS>
- 2.<IMAGE>\nIdentify the intervention phase shown in this image: <OPTIONS>
- 3.<IMAGE>\nWhat is the intervention-related state depicted in this image? <OPTIONS>
- 4.<IMAGE>\nSelect the intervention phase that best describes this image: <OPTIONS>
- 5.<IMAGE>\nWhich category best describes the intervention phase in this image: <OPTIONS>

Conversation example:



- **Human:** What is the intervention-related state depicted in this image? <A> dyed resection margin, resected polyp, <C> resection margin, <D> dyed lifted polyp
- **Machine:** <D> dyed lifted polyp

MUT#5 ♦ Imaging Quality Assessment

Task definition: Evaluate the quality of a colonoscopy image by identifying any relevant exposure conditions.

Task type: Procedural quality control

Evaluation metrics: Accuracy

Category: CLS#72 ~ CLS#74

Dataset origin: DATA#31

Number of images: Training 2741 | Validation 132 | Testing 1,140

Number of VQA entries: Training 2741 | Validation 132 | Testing 1,140

Task instruction templates:

- 1.<IMAGE>\nCheck the quality of the provided colonoscopy image and confirm any issues. <OPTIONS>
- 2.<IMAGE>\nAssess whether the given colonoscopy image meets quality standards. <OPTIONS>
- 3.<IMAGE>\nEvaluate the provided colonoscopy image for any quality concerns. <OPTIONS>
- 4.<IMAGE>\nDetermine if the given colonoscopy image is of acceptable quality. <OPTIONS>
- 5.<IMAGE>\nReview the colonoscopy image to identify issues or confirm good quality. <OPTIONS>

Conversation example:



- ▶ **Human:** Determine if the given colonoscopy image is of acceptable quality. <A> normal exposure, overexposed, <C> underexposed
- ▶ **Machine:** <A> normal exposure

MUT#6 ♦ Imaging Modality Classification

Task definition: Distinguish the optical imaging modality used in colonoscopy.

Task type: Procedural quality control

Evaluation metrics: Accuracy

Category: CLS#42 ~ CLS#46

Dataset origin: DATA#26 ~ DATA#28

Number of images: Training 11,781 | Validation 2,498 | Testing 5,110

Number of VQA entries: Training 11,781 | Validation 2,498 | Testing 5,110

Task instruction templates:

1.<IMAGE>\nWhat is the imaging method shown in this image? <OPTIONS>

2.<IMAGE>\nWhich imaging technique does this image use? <OPTIONS>

3.<IMAGE>\nHow was the image captured? <OPTIONS>

4.<IMAGE>\nIdentify the imaging technique used for this image. <OPTIONS>

5.<IMAGE>\nWhat type of imaging is depicted in this image? <OPTIONS>

Conversation example:



► **Human:** What is the imaging method shown in this image? <A> Linked Color Imaging (LCI), Blue Light Imaging (BLI), <C> Narrow Band Imaging (NBI), <D> White Light Imaging (WLI), <E> Flexible Imaging Color Enhancement (FICE)

► **Machine:** <C> Narrow Band Imaging (NBI)

MUT#7 ♦ Instrument Recognition

Task definition: Detect surgical instruments in colonoscopy images to ensure proper tracking and operational safety.

Task type: Safety monitoring

Evaluation metrics: Accuracy

Category: CLS#47 ~ CLS#49, CLS#75

Dataset origin: DATA#11 ~ DATA#13, DATA#29

Number of images: Training 2,468 | Validation 260 | Testing 1,006

Number of VQA entries: Training 2,468 | Validation 260 | Testing 1,006

Task instruction templates:

- 1.<IMAGE>\nDo you see any surgical instruments in this colonoscopy image? Answer me with Yes or No.
- 2.<IMAGE>\nAre there any medical tools visible in the colonoscopy image? Answer me with Yes or No.
- 3.<IMAGE>\nCan you identify any surgical equipment in the colonoscopy image? Answer me with Yes or No.
- 4.<IMAGE>\nIs there any surgical tools present in this colonoscopy image? Answer me with Yes or No.
- 5.<IMAGE>\nAre there any surgical tools shown in the colonoscopy image? Answer me with Yes or No.

Conversation example:



► **Human:** Do you see any surgical instruments in this colonoscopy image? Answer me with Yes or No.

► **Machine:** Yes

MUT#8 ♦ Bleeding Warning

Task definition: Detect surgical instruments in colonoscopy images to ensure proper tracking and operational safety.

Task type: Safety monitoring

Evaluation metrics: Accuracy

Category: CLS#21 ~ CLS#24, CLS#76

Dataset origin: DATA#12 ~ DATA#17

Number of images: Training 2,596 | Validation 258 | Testing 1,224

Number of VQA entries: Training 2,596 | Validation 258 | Testing 1,224

Task instruction templates:

- 1.<IMAGE>\nCan you spot any areas of bleeding in the colonoscopy image? Answer me with Yes or No.
- 2.<IMAGE>\nIs there any visible bleeding in this colonoscopy image? Answer me with Yes or No.
- 3.<IMAGE>\nAre there regions showing blood in the colonoscopy image? Answer me with Yes or No.
- 4.<IMAGE>\nDo you see any bleeding in the colonoscopy image? Answer me with Yes or No.
- 5.<IMAGE>\nIs there any sign of bleeding in this colonoscopy image? Answer me with Yes or No.

Conversation example:



► **Human:** Are there regions showing blood in the colonoscopy image? Answer me with Yes or No.

► **Machine:** Yes

MUT#9 ♦ Lesion diagnosis (yes-or-no mode)

Task definition: Determine the presence of specific abnormalities to identify categories in colonoscopy images, aiding diagnosis and treatment.

Task type: Lesion diagnosis

Evaluation metrics: Accuracy

Category: CLS#1 ~ CLS#41

Dataset origin: DATA#1 ~ DATA#26

Number of images: Training 79,873 | Validation 7,897 | Testing 45,246

Number of VQA entries: Training 159,746 | Validation 15,794 | Testing 90,492

Task instruction templates:

- 1.<IMAGE>\nDoes the image contain a {category}? Answer me with Yes or No.
- 2.<IMAGE>\nIs there a {category} region visible in the image? Answer me with Yes or No.
- 3.<IMAGE>\nDoes the image show any {category} areas? Answer me with Yes or No.
- 4.<IMAGE>\nCan you identify any {category} in the image? Answer me with Yes or No.
- 5.<IMAGE>\nDoes the image display any {category} area? Answer me with Yes or No.

Conversation example:



- **Human:** Does the image show any angiectasia areas? Answer me with Yes or No.
- **Machine:** Yes

MUT#10 ♦ Lesion diagnosis (single-choice mode)

Task definition: Identify abnormality types in colonoscopy images, selecting from two to ten predefined options for diagnostic and treatment purposes.

Task type: Lesion diagnosis

Evaluation metrics: Accuracy

Category: CLS#1 ~ CLS#41

Dataset origin: DATA#1 ~ DATA#26

Number of images: Training 79,873 | Validation 7,897 | Testing 45,246

Number of VQA entries: Training 79,873 | Validation 7,897 | Testing 45,246

Task instruction templates:

- 1.<IMAGE>\nChoose the correct category for this colonoscopy image from the following options.
<OPTIONS>
- 2.<IMAGE>\nWhich category best describes the colonoscopy image? <OPTIONS>
- 3.<IMAGE>\nSelect the most fitting category based on this colonoscopy image. <OPTIONS>
- 4.<IMAGE>\nFrom the provided options, identify the category that matches the colonoscopy image content
<OPTIONS>
- 5.<IMAGE>\nBased on the colonoscopy image, pick the appropriate category to describe. <OPTIONS>

Conversation example:



- **Human:** Select the most fitting category based on this colonoscopy image. <A> blood in lumen, polyp, <C> bleeding, <D> vascular anomalies
- **Machine:** polyp

MUT#11 ♦ Lesion diagnosis (open-vocabulary mode)

Task definition: Identify abnormality categories in colonoscopy images for early diagnosis and treatment planning, unrestricted by predefined categories.

Task type: Lesion diagnosis

Evaluation metrics: Accuracy

Category: CLS#1 ~ CLS#41

Dataset origin: DATA#1 ~ DATA#26

Number of images: Training 79,873 | Validation 7,897 | Testing 45,246

Number of VQA entries: Training 79,873 | Validation 7,897 | Testing 45,246

Task instruction templates:

- 1.<IMAGE>\nCategorize the object in the given colonoscopy image.
- 2.<IMAGE>\nDetermine the object's category in this colonoscopy image.
- 3.<IMAGE>\nIdentify the category of the object in the provided colonoscopy image.
- 4.<IMAGE>\nClassify the object's category in the given colonoscopy image.
- 5.<IMAGE>\nAssign the object to its corresponding category in the colonoscopy image.

Conversation example:



- **Human:** Classify the object's category in the given colonoscopy image
- **Machine:** high grade adenoma

MUT#12 ♦ Referring Expression Generation

Task definition: Generate the category name of an object in a specified region of the image.

Task type: Lesion diagnosis

Evaluation metrics: Accuracy

Category: CLS#11 ~ CLS#24, CLS#27 ~ CLS#29, CLS#41, CLS#47 ~ CLS#50

Dataset origin: DATA#1 ~ DATA#9, DATA#11, DATA#12, DATA#14, DATA#15, DATA#17 ~ DATA#20, DATA#29, DATA#30

Number of images: Training 79,611 | Validation 6,426 | Testing 50,865

Number of VQA entries: Training 79,611 | Validation 6,426 | Testing 50,865

Task instruction templates:

- 1.<IMAGE>\nWhat category does object coordinates belong to?
- 2.<IMAGE>\nCan you tell me the category of object coordinates?
- 3.<IMAGE>\nCould you provide the category for object coordinates?
- 4.<IMAGE>\nPlease specify the category of object coordinates.
- 5.<IMAGE>\nWhat is the category for object coordinates.

Conversation example:



- **Human:** What is the category for {<40><155><263><279>}.
- **Machine:** polyp

MUT#13 ♦ Referring Expression Comprehension

Task definition: Locate and provide the bounding box for a specified category in the colonoscopy image.

Task type: Lesion diagnosis

Evaluation metrics: IoU

Category: CLS#11 ~ CLS#24, CLS#27 ~ CLS#29, CLS#41, CLS#47 ~ CLS#50

Dataset origin: DATA#1 ~ DATA#9, DATA#11, DATA#12, DATA#14, DATA#15, DATA#17 ~ DATA#20, DATA#29, DATA#30

Number of images: Training 79,611 | Validation 6,426 | Testing 50,865

Number of VQA entries: Training 79,611 | Validation 6,426 | Testing 50,865

Task instruction templates:

- 1.<IMAGE>\nWhere is the location of object category?
- 2.<IMAGE>\nCould you give the position of object category?
- 3.<IMAGE>\nWhere is object category located?
- 4.<IMAGE>\nCould you specify the location of object category?
- 5.<IMAGE>\nPlease specify the coordinates of object category

Conversation example:



- **Human:** Where is the location of polyp?
- **Machine:** {<40><13><133><266>}

MUT#14 ♦ Early Cancer Grading (NICE criteria)

Task definition: Assess polyps with NICE classification systems.

Task type: Disease grading

Evaluation metrics: Accuracy

Category: CLS#55 ~ CLS#57

Dataset origin: DATA#7

Number of images: Training 2,083 | Validation 872 | Testing 310

Number of VQA entries: Training 2,083 | Validation 872 | Testing 310

Task instruction templates:

1.<IMAGE>\nPlease classify the polyp in this colonoscopy image using the NICE classification system.

<OPTIONS>

2.<IMAGE>\nBased on the features observed in this colonoscopy image, classify the polyp according to the NICE criteria. <OPTIONS>

3.<IMAGE>\nClassify the type of polyp in this colonoscopy image based on its characteristics using the NICE classification system. <OPTIONS>

4.<IMAGE>\nAssess the polyp in this colonoscopy image and classify its type according to the NICE classification. <OPTIONS>

5.<IMAGE>\nPlease evaluate the type of polyp in this colonoscopy image using the NICE classification system. <OPTIONS>

Conversation example:



► **Human:** Please classify the polyp in this colonoscopy image using the NICE classification system. <A> Type 1 (characteristic for hyperplastic polyp), Type 3 (characteristic for malignancy), <C> Type 2 (characteristic for adenoma)

► **Machine:** <A> Type 1 (characteristic for hyperplastic polyp)

MUT#15 ♦ Early Cancer Grading (PARIS criteria)

Task definition: Assess polyps with PARIS classification systems.

Task type: Disease grading

Evaluation metrics: Accuracy

Category: CLS#58 ~ CLS#61

Dataset origin: DATA#16

Number of images: Training 19,544 | Validation 0 | Testing 29,592

Number of VQA entries: Training 19,544 | Validation 0 | Testing 29,592

Task instruction templates:

1.<IMAGE>\nBased on the PARIS classification criteria, what is the type of the polyp in this colonoscopy image? <OPTIONS>

2.<IMAGE>\nPlease classify the type of polyp in this colonoscopy image according to the PARIS classification system. <OPTIONS>

3.<IMAGE>\nBased on the features observed in this colonoscopy image, classify the polyp type using the PARIS criteria. <OPTIONS>

4.<IMAGE>\nIn this colonoscopy image, identify the type of polyp based on the PARIS classification system. <OPTIONS>

5.<IMAGE>\nClassify the polyp in this colonoscopy image using the PARIS classification system based on its appearance. <OPTIONS>

Conversation example:



► **Human:** In this colonoscopy image, identify the type of polyp based on the PARIS classification system.
<A> Sessile (Is), Subpedunculated (Isp), <C> Slightly elevated (IIa), <D> Pedunculated (Ip)

► **Machine:** <A> Sessile (Is)

MUT#16 ♦ Polyp Sizing

Task definition: Evaluate polyp size in colonoscopy images based on specified ranges.

Task type: Disease grading

Evaluation metrics: Accuracy

Category: CLS#63 ~ CLS#66

Dataset origin: DATA#7, DATA#16

Number of images: Training 21,512 | Validation 706 | Testing 29,775

Number of VQA entries: Training 21,512 | Validation 706 | Testing 29,775

Task instruction templates:

1.<IMAGE>\nAnalyze the colonoscopy image to estimate the size of the polyp and classify it into the appropriate size range. <OPTIONS>

2.<IMAGE>\nAssess the colonoscopy image to gauge the polyp's size and classify it within the given size options. <OPTIONS>

3.<IMAGE>\nAnalyze this colonoscopy image to estimate the size of the polyp and sort it into the corresponding size range. <OPTIONS>

4.<IMAGE>\nEvaluate the provided colonoscopy image to determine the approximate size range of the polyp. <OPTIONS>

5.<IMAGE>\nInspect the colonoscopy image to measure the polyp's size and label it within the appropriate size range. <OPTIONS>

Conversation example:



- **Human:** Assess the colonoscopy image to gauge the polyp's size and classify it within the given size options.
<A> polyp \geq 30mm, 20mm \leq polyp < 30mm, <C> 0mm < polyp < 6mm, <D> 6mm \leq polyp < 20mm
- **Machine:** <C> 0mm < polyp < 6mm

MUT#17 ♦ Ulcerative Colitis Grading

Task definition: Grade the severity of ulcerative colitis using the Mayo score system.

Task type: Disease grading

Evaluation metrics: Accuracy

Category: CLS#34 ~ CLS#40

Dataset origin: DATA#10, DATA#22 , DATA#23

Number of images: Training 10,253 | Validation 110 | Testing 2,015

Number of VQA entries: Training 10,253 | Validation 110 | Testing 2,015

Task instruction templates:

- 1.<IMAGE>\nHow would you classify the severity of the ulcer in the image? <OPTIONS>
- 2.<IMAGE>\nWhat stage does the ulcer appear to be in this picture? <OPTIONS>
- 3.<IMAGE>\nCan you determine the grade of the ulcer shown in the photo? <OPTIONS>
- 4.<IMAGE>\nWhat level of severity is the ulcer in the image? <OPTIONS>
- 5.<IMAGE>\nHow advanced is the ulcer depicted in this picture? <OPTIONS>

Conversation example:



- **Human:** How advanced is the ulcer depicted in this picture? <A> ulcerative colitis grade 1-2, ulcerative colitis grade 0-1, <C> ulcerative colitis grade 2-3, <D> ulcerative colitis grade 3, <E> ulcerative colitis grade 1, <F> ulcerative colitis grade 0, <G> ulcerative colitis grade 2
- **Machine:** ulcerative colitis grade 0-1

Task definition: Generate a detailed textual description of the colonoscopy image.

Task type: Captioning

Category source: CLS#1 ~ CLS#11, CLS#15, CLS#22 ~ CLS#30, CLS#32 ~ CLS#36, CLS#38 ~ CLS#41, CLS#47 ~ CLS#54, CLS#66 ~ CLS#71

Dataset source: DATA#2 ~ DATA#4, DATA#20, DATA#21, DATA#25, DATA#26, DATA#29, DATA#32

Number of images: Training 70,745 | Validation 8,996 | Testing 44,157

Number of VQA entries: Training 70,745 | Validation 8,996 | Testing 44,157

Task instruction templates:

1. <IMAGE>\nDescribe what you see in the image.
2. <IMAGE>\nInterpret what the image shows.
3. <IMAGE>\nDetail the visual elements in the image.
4. <IMAGE>\nExplain the image's visuals thoroughly.
5. <IMAGE>\nOffer a thorough explanation of the image.

Conversation example:



► **Human:** Explain the image's visuals thoroughly.

► **Machine:** The image displays a close-up view of a colorectal polyp as seen during an endoscopic procedure. The polyp appears to be sessile with a slightly lobulated surface, which can be characteristic of adenomatous polyps, a common type of polyp that can be pre-malignant. The coloration of the polyp is somewhat homogenous, lacking the reddened or ulcerated appearance that might suggest active inflammation or more concerning pathology. Surrounding mucosa seems to be intact and normal in color, suggesting the polyp is an isolated finding without immediately visible widespread mucosal abnormalities. Given the visual characteristics of the polyp, further medical evaluation, such as a biopsy, may be warranted to determine its histological nature and potential for malignancy