DATASET SCOPE

Foreign-object annotations for HEICO training frames

OReNA FOCUS FRAME | source snapshot v14, 7 September 2026

Tristan Kirscher[1,2] **Lalith Sharan**[3,4] **Andrea Naclerio**[3,4] **Luc Vedrenne**[3,4] **Nicolas Padoy**[3,4]

[1]ICube Laboratory, CNRS UMR-7357, University of Strasbourg, Strasbourg, France
[2]CLCC Institut Strauss, Strasbourg, France
[3]CAMMA, ICube, UMR7357, University of Strasbourg, CNRS, INSERM, Strasbourg, France
[4]Institute of Image-Guided Surgery, IHU Strasbourg, Strasbourg, France

1,274	20	2,505	4,228
annotated frames	train videos	box instances	flag-filtered QA

This annotation release covers the HEICO training subset of the v14 snapshot. It provides reviewed bounding boxes, structured per-frame facts and question–answer (QA) pairs derived from those annotations. Original images, videos, model weights and LapChole data are excluded.

Object coverage

The counts below are annotated instances across frames, not distinct physical objects tracked across videos. The taxonomy defines ten foreign-object categories; seven are observed in this subset.

Category	Boxes	Relative count
Sponge	291	
Clip	896	
Specimen Bag	74	
Silicone Loop	410	
External Drain	524	
Needle	256	
Gallstone	0	No examples
Specimen	54	
Mesh	0	No examples
AHA (provisional)	0	No examples

AHA: Absorbable Hemostatic Agent, a provisional taxonomy category. Zero examples indicate absence from this release, not clinical absence. Bar lengths share the same linear scale.

What the release represents

The source records describe human acceptance, correction or drawing of boxes after detector/SAM 2 assistance. They designate 834 frames as label-assisted and 440 as image-only annotation. All exported records are marked completed and exhaustive in that workflow; these statuses are not independent measurements of annotation accuracy.

The sample is selected training material. Detector-assisted selection and temporal propagation produce dependent observations. Keep each video's frames, crops and derived QA together when constructing downstream splits. The package is not an unbiased calibration or test set, or a representative estimate of object prevalence.

DOI: 10.5281/zenodo.22751012. Annotations and this note: CC BY 4.0. Source images are excluded and retain their own terms. Source attribution appears on page 3.

FORMATS AND LOADING

From annotation to model input

Files and their roles

File	Contents and use
`annotations.jsonl`	Structured annotation records; identities, free-text notes and annotation timestamps omitted.
`coco.json`	2,505 pixel XYWH boxes with image dimensions and preserved category/instance IDs.
`frame_facts.jsonl`	1,274 per-frame records with counts and boxes.
`dino_crops.jsonl`	2,505 positive crop specifications and 15 `frame_negative` records; no crop pixels.
`source_image_manifest.jsonl`	One-to-one mapping between task IDs, export aliases, original relative paths and expected image hashes.
`quality_flags.jsonl`	86 flagged frames; their task IDs define the existing QA exclusion filter.
`annotation_taxonomy.json`	Category IDs, names, definitions and controlled attribute values.
`PACKAGE_MANIFEST.json`	Expected file sizes and SHA-256 hashes for the packaged files.

Resolve source images explicitly

`image_file_name` and `image_path` are export aliases, not original source paths. COCO stores the same alias in `file_name`. Look up the alias or `task_id` in `source_image_manifest.jsonl`; join `source_relative_path` to your authorized HEICO root containing `frames/`, then verify `image_sha256`. Do not flatten paths to basenames.

All 1,274 aliases and original relative paths are unique in this package. Hashes were checked for consistency against source metadata; they were not recomputed from original image bytes during release preparation. No dataset downloader is included.

Choose a QA input once

Input	Rows
Raw canonical QA: `synthetic_vqa.jsonl`	4,586
Canonical QA after the existing flag filter	4,228
Optional text-QA union after the flag filter	7,013

The latter two files are `synthetic_vqa_clean.jsonl` and `augmented_vqa_clean.jsonl` under `optional_vlm_augmentation/`. The optional union duplicates its four constituent family files: use the union *or* those files, never both. Family names containing `balanced` describe generation before HEICO selection; the public subsets are not exactly balanced.

Detection filtering

COCO, structured annotations and DINO retain flagged frames. Applying the QA filter to detection training requires removing the flagged task IDs and, for COCO, their images and dependent annotations. There are 212 DINO records on flagged frames. Empty-frame records are frame-level negatives, not annotated instrument/tissue hard-negative boxes.

Interpreting the annotations

What the existing flags mean

Excluding 86 flagged frames retains 1,188 frames and 4,228 canonical QA rows. The flags record specified annotation/source-hint discrepancies, declared class ambiguity and conflicting source answers. They are not a comprehensive visual-quality screen or independent clinical adjudication.

Recorded attribute or status	Count
Boxes with `visibility=uncertain`	1,012 / 2,505
Uncertain-visibility boxes after the frame flag filter	1,010
Sponge boxes with an uncertain state	180 / 291
Boxes marked `truncated=true`	0
Annotated empty frames	15

These are recorded fields, not probabilities of annotation error. Retained frames may contain glare, blur, blood or occlusion. All truncation fields being false does not independently establish absence of truncation. Distinguish default/unfilled attributes from actual visual judgments before using them as a validated attribute benchmark.

Counting and class boundaries

Historical auditing found Clip undercounts in an earlier subset: 10 disagreements among 118 comparable historical frames. Count-consistency checks introduced for later additions were not applied uniformly to inherited frames. Residual undercounting has not been quantified for this release; the historical fraction is not its current error rate.

The taxonomy distinguishes placed clips from clips loaded in an applicator, detached specimens from attached anatomy, and visible needles from suture alone. Difficult single-frame cases require image-level adjudication. *Foreign object* is a challenge convention, not a clinical judgment that an object is unwanted, forgotten or retained. Empty-frame and zero/absence QA describe the visible field, not an entire operation.

Preserved information and validation limits

Structured labels, annotation protocols, visibility fields, image identifiers and hashes are retained. Annotator identities, free-text notes, annotation timestamps and source question IDs are omitted. There were 755 source annotations with nonempty notes; loss of scientific information from those notes has not been independently adjudicated.

Geometry, cross-file references, QA consistency with the recorded boxes and package integrity were checked. Images were not independently clinically reviewed during release preparation. The exact use of this snapshot by a submitted model remains unconfirmed. Source snapshot and export identity are recorded in `release_provenance.json`; package hashes are in the file manifest.

Source attribution

Maier-Hein, L., Wagner, M., Ross, T. et al. *Heidelberg colorectal data set for surgical data science in the sensor operating room*. Scientific Data **8**, 101 (2021). doi:10.1038/s41597-021-00882-2. This is the source video dataset reference, not the citation for these additional annotations.

OReNA FOCUS FRAME challenge: https://or-arena.org/. HEICO FOCUS VQA source: https://huggingface.co/datasets/orena-dkfz/heico-focus-vqa. Obtain source images and applicable access terms from the dataset owners.