

General (26.6%)

- Image Textualization (Filt.) (24.9 K)
- ShareGPT-4V (COCO) (12.5 K)
- ALLaVA-Instr. (V-FLAN-4V) (5.0 K)
- LLaVA-Wild 12K (3.8 K)
- ScienceQA (Cauldron/LLaVA) (1.2 K)

- TallyQA (Cauldron/LLaVA) (24.7 K)
- ALLaVA-Instr. (LAION-4V) (12.5 K)
- LLaVAR (GPT-4) (4.9 K)
- Visual7W (Cauldron/LLaVA) (3.6 K)
- VSR (Cauldron/LLaVA) (538)

- CLEVR (Cauldron/LLaVA) (17.5 K)
- LLaVA-Wild 39K (9.8 K)
- ST-VQA (Cauldron/LLaVA) (4.3 K)
- VisText (Cauldron) (2.5 K)
- InterGPS (Cauldron/LLaVA) (318)

- ShareGPT-40 (14.3 K)
- ShareGPT-4V (LLaVA) (7.5 K)
- AOK-VQA (Cauldron/LLaVA) (4.1 K)
- VQARAD (Cauldron/LLaVA) (77)

Hateful Memes (Cauldron/LLaVA) (2.1 K)

Doc/Chart/Screen (28.3%)

- DVQA (Cauldron/LLaVA) (50.0 K)
- RoBUT-WTQ (Cauldron/LLaVA) (9.6 K)
- UReader-IE (4.3 K)
- RoBUT-SQA (Cauldron) (2.1 K)
- HiTab (Cauldron/LLaVA) (623)
- LRV-Chart (444)

- FigureQA (Cauldron/LLaVA) (25.0 K)
- TQA (Cauldron/LLaVA) (6.8 K)
- Screen2Words (Cauldron) (3.9 K)
- AI2D (GPT-4V) (1.2 K)
- AI2D (Cauldron/LLaVA) (607)
- Diagram2Text (Cauldron) (73)

- UReader-Cap (22.9 K)
- Chart2Text (Cauldron) (6.7 K)
- AI2D (InternVL) (3.1 K)
- Infographic-VQA (1.1 K)
- Infographic-VQA (LLaVA) (528)

- RoBUT-WikiSQL (Cauldron) (18.7 K)
- ChartQA (Cauldron/LLaVA) (4.6 K)
- LRV-Normal (Filt.) (2.6 K)
- VisualMRC (Cauldron) (755)
- Infographic (GPT-4V) (495)

Math/Reasoning (27.5%)

- MAVIS RuleGeo (25.0 K)
- Raven (Cauldron) (10.5 K)
- IconQA (Cauldron/LLaVA) (6.8 K)
- GeoQA+ (MathV360K) (4.3 K)
- GeoMVerse (Cauldron) (2.3 K)
- CLEVR-Math (MathV360K) (1.3 K)
- GEOS (MathV360K) (124)

- MAVIS MetaGen (21.8 K)
- MapQA (Cauldron/LLaVA) (9.4 K)
- IconQA (MathV360K) (5.6 K)
- UniGeo (MathV360K) (3.0 K)
- Super-CLEVR (MathV360K) (2.2 K)
- MapQA (MathV360K) (1.3 K)

- Geo170K-QA (17.0 K)
- PMC-VQA (MathV360K) (9.0 K)
- TabMWP (MathV360K) (5.6 K)
- TQA (MathV360K) (2.5 K)
- VizWiz (MathV360K) (1.7 K)
- VIZWIZ (MatilV300K) (1.7 F
- VQA-RAD (MathV360K) (531)

- Geo170K-Align (15.1 K)
- MathQA (7.5 K)
- FigureQA (MathV360K) (4.4 K)
- Geometry3K (MathV360K) (2.4 K)
- VQA-AS (MathV360K) (1.5 K)
- Geo3K (522)

General OCR (17.6%)

- K-12 Printing (64.2 K)
- Rendered Text (Cauldron) (2.5 K)
- IIIT-5K (497)

- HME-100K (18.6 K)
- WebSight (Cauldron) (2.5 K)

- TextOCR (GPT-4V) (6.3 K)
- Chrome Writing (2.2 K)

- TextCaps (5.5 K)
- IAM (Cauldron) (1.4 K)