Dataset	# images	# QA pairs	# tokens	% mix
Captioning				
ShareGPT-4o [1]	57,259	57,259	39,696,010	13.03%
LNarratives LocalizedNarratives	507,444	507,444	21,328,731	1.40%
TextCaps textcaps	21,953	21,953	389,658	1.28%
VisText VisText	7,057	9,969	1,245,485	1.23%
IIW-400 IIW-400	400	400	103,024	0.68%
Screen2Words screen2words	15,730	15,743	143,103	0.23%
Real-world visual question answering				
LNQA [2]	302,780	1,520,942	21,107,241	3.46%
VQAv2 VQAv2	82,772	443,757	1,595,929	2.10%
COCO-QA CocoQA	46,287	78,736	286,982	0.94%
Visual7W Visual7w	14,366	69,817	279,268	0.92%
OK-VQA okvqa	8,998	9,009	38,853	0.26%
VSR VSR	2,157	3,354	10,062	0.13%
OCR, document understanding, text transcription				
Docmatix[3] (ours)	1,273,215	9,488,888	392,302,612	10.31%
RenderedText[4]	999,000	999,000	27,207,774	7.15%
DocVQA DocVQA	10,189	39,463	337,829	2.22%
TextVQA textvqa	21,953	34,602	181,918	1.19%
Cord-v2 [5]	800	800	178,388	1.17%
ST-VQA STVQA	17,247	23,121	127,846	0.84%
OCR-VQA OCR-VQA	165,746	801,579	6,073,824	0.60%
VisualMRC VisualMRC	3,027	11,988	168,828	0.55%
IAM IAM	5,663	5,663	144,216	0.47%
InfoVQA InfographicVQA	2,118	10,074	61,048	0.40%
Diagram image-to-text[6]	300	300	22,196	0.07%
Chart/figure understanding				
Chart2Text Chart2Text	26,985	30,242	2,852,827	4.38%
DVQA DVQA	200,000	2,325,316	8,346,234	4.27%
ChartQA ChartQA	18,271	28,299	185,835	1.90%
PlotQA PlotQA	157,070	20,249,479	8478299.278	0.65%
FigureQA FigureQA	100,000	1,327,368	3,982,104	0.61%
MapQA MapQA	37,417	483,416	6,470,485	0.33%
Table understanding				
TabMWP TabMWP	22,729	23,059	1,948,166	1.60%
TAT-QA TAT-QA	2,199	13,215	283,776	1.40%
HiTab Hitab	2,500	7,782	351,299	1.15%
MultiHiertt Multihiertt	7,619	7,830	267,615	0.88%
FinQA FinQA	5,276	6,251	242,561	0.64%
WikiSQL WikiSQL	74,989	86,202	9,680,673	0.64%
SQA SQA	8,514	34,141	1,894,824	0.62%
TQA TQA	1,496	6,501	26,004	0.34%
WTQ WTQ	38,246	44,096	6,677,013	0.33%
Reasoning, logic, maths, geometry				
Geo170K G-llava-Geo170K	9,067	177,457	17,971,088	2.95%
GeomVerse GeomVerse	9,303	9,339	2,489,459	2.45%
CLEVR-Math CLEVR-Math	70,000	788,650	3,184,656	2.09%
CLEVR CLEVR	70,000	699,989	2,396,781	0.79%
A-OKVQA A-OKVQA	16,539	17,056	236,492	0.78%
IconQA IconQA	27,315	29,859	112,969	0.74%
AI2D AI2D	3,099	9,708	38,832	0.51%
NLVR2 NLVR2	50,426	86,373	259,119	0.43%
RAVEN RAVEN	42,000	42,000	105,081	0.43%
TallyQA TallyQA	98,680	183,986	738,254	0.36%
Spot the diff SpotTheDiff	8,566	9,524	221,477	0.36%
GSD MIMIC-IT-General-Scene-Difference	70,939	141,869	4,637,229	0.30%
ScienceQA ScienceQA	4,985	6,218	24,872	0.16%
Inter-GPs Inter-GPS	1,451	2,101	8,404	0.11%
HatefulMemes hatefulmeme	8,500	8,500	25,500	0.08%
Screenshot to code				
WebSight WebSight	500,000	500,000	276,743,299	0.91%
DaTikz DaTikz	47,974	48,296	59,556,252	0.02%
Text-only general instructions, math problems, arithmetic calculations				
OpenHermes-2.5 OpenHermes	0	1,006,223	248,553,747	8.16%
MetaMathQA MetaMathQA	0	395,000	74,328,255	2.44%
AtlasMathSets[7]	0	17,807,579	455,411,624	2.24%
MathInstruct MathInstruct	0	261,781	45,393,559	1.49%
OrcaMath Orca-Math	0	200,031	63,780,702	1.05%
Goat Goat	0	1,746,300	167,695,693	0.55%
LIMA LIMA	0	1,052	633,867	0.52%
Dolly Dolly	0	14,972	1,329,999	0.44%
CamelAIMath CamelAIMath	0	49,744	21,873,629	0.04%