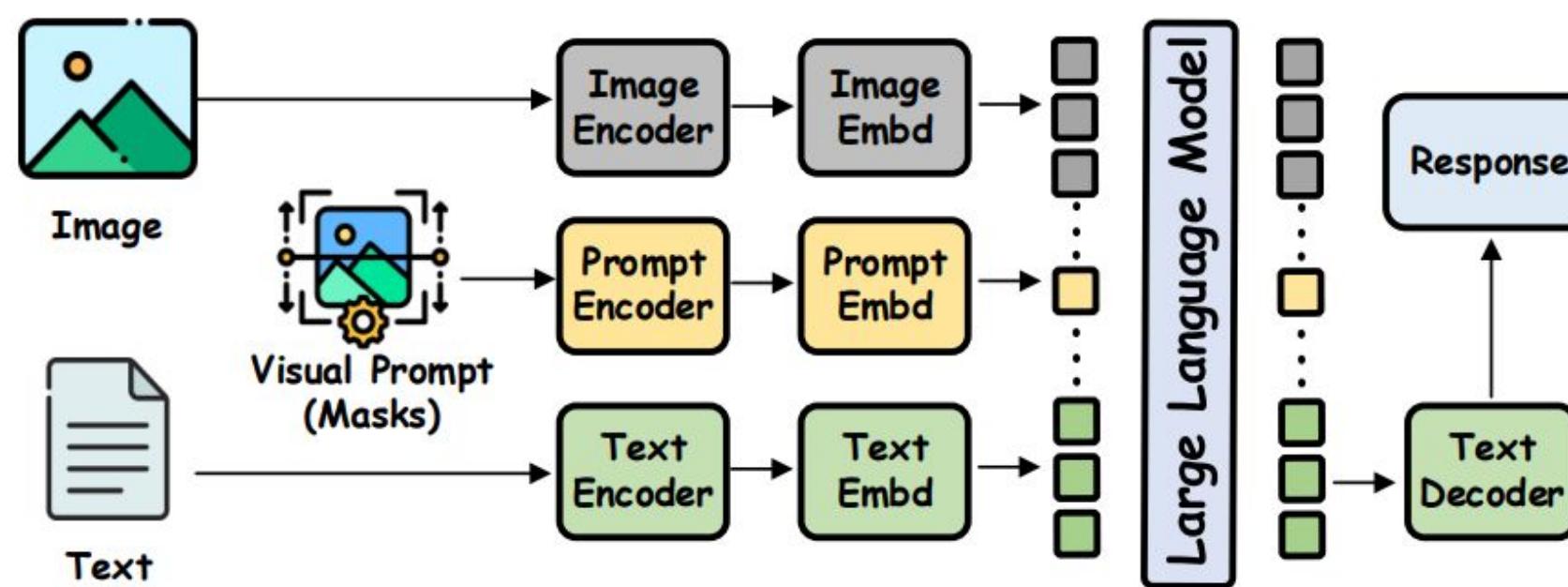


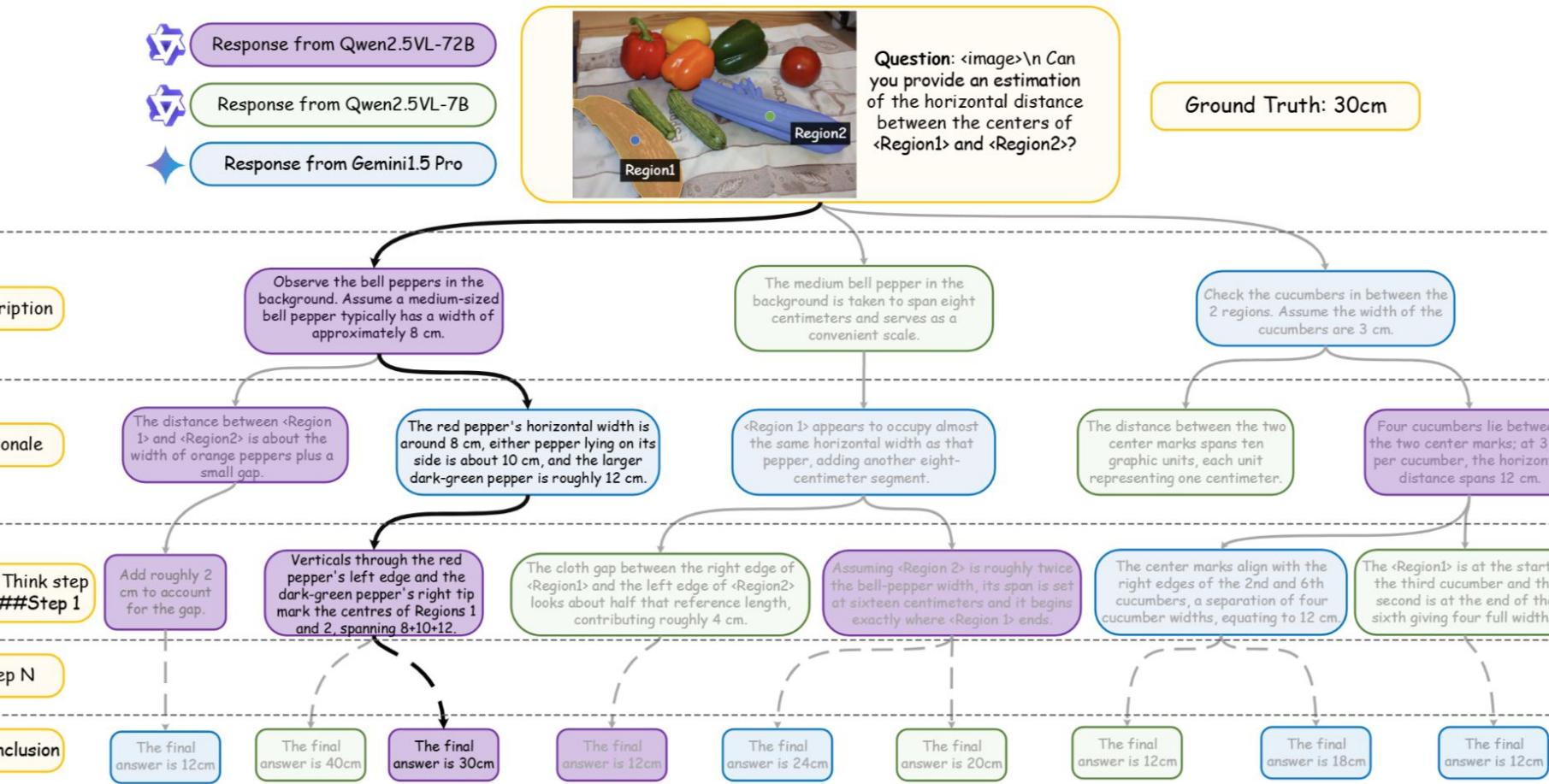
Overview

- Current MLLMs struggle with fine-grained visual understanding and coherent step-by-step reasoning.
- We propose a three-stage pipeline: generate LongCoT with **M3CTS**, construct preference pairs, apply a novel **fine-grained Preference Optimization (fDPO)** method that employs segment-specific learning updates tailored for descriptive grounding and logical reasoning

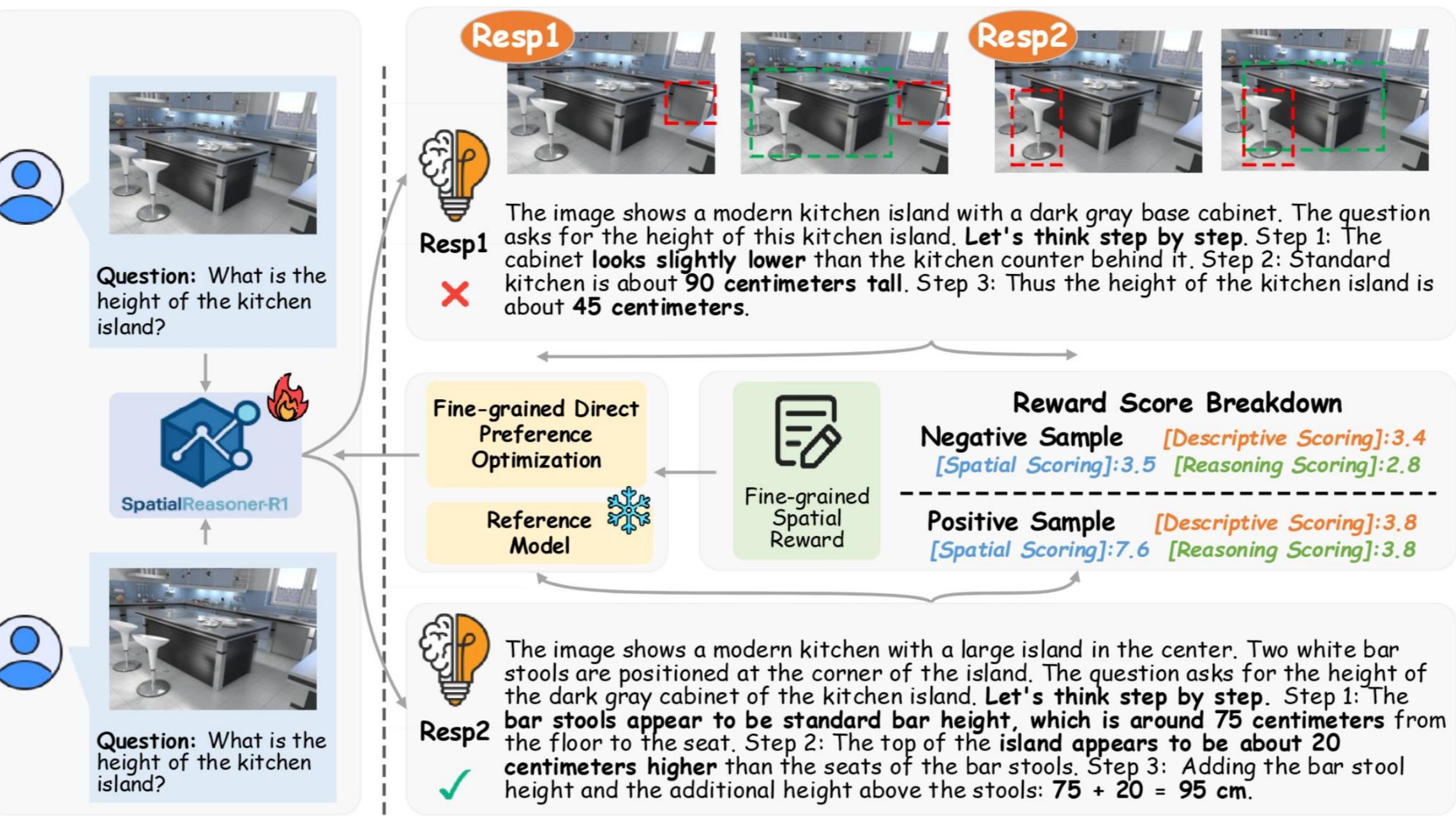
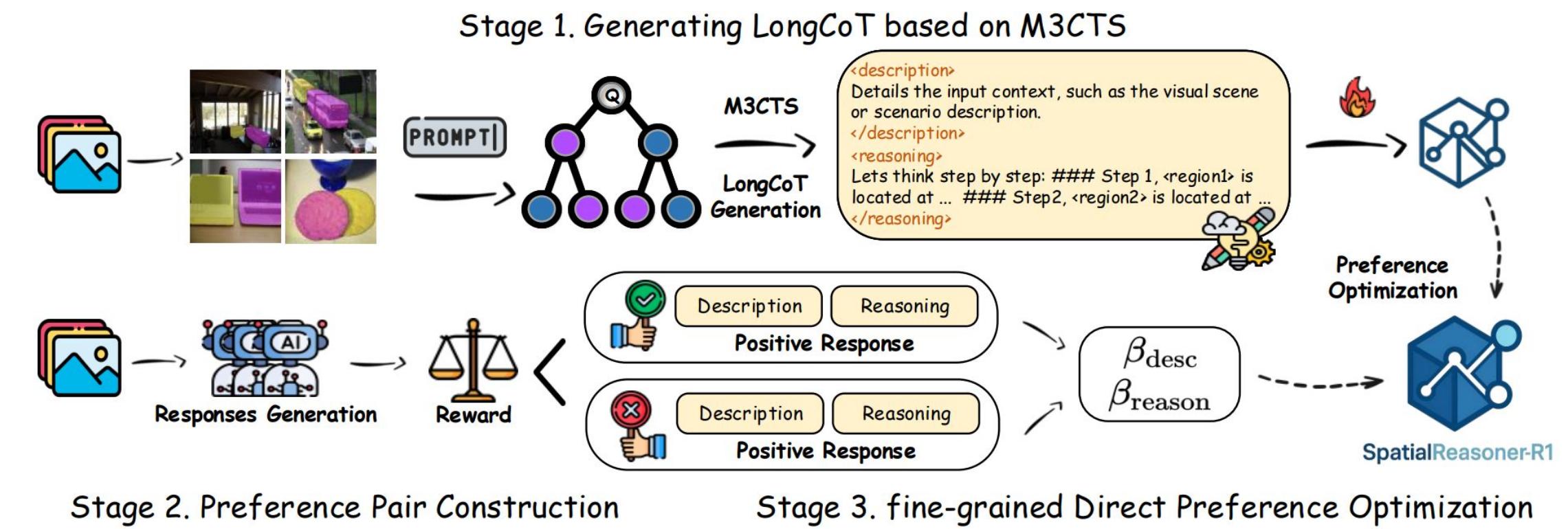
Architecture



Dataset Pipeline



Method



Experiment

	Below/Above	Left/Right	Big/Small	Tall/Short	Wide/Thin	Behind/Front	Qual Acc
Gemini 2.0 Flash [21]	58.33	68.57	16.98	50.00	15.38	53.63	44.29
Llama 4 Maverick [47]	54.17	61.90	33.02	50.89	25.96	55.45	47.18
Gemini 1.5 Pro [63]	85.83	56.19	58.49	71.42	55.76	60.00	65.14
ChatGPT-4o [49]	87.50	80.00	53.77	63.39	51.92	60.90	66.67
SpatialBot-3B [6]	52.50	62.86	57.54	49.11	49.04	62.73	55.56
SpaceThinker Qwen2.5VL-3B [4]	89.16	63.81	76.41	56.25	56.73	70.91	69.25
InternVL2.5-78B [12]	94.16	94.28	64.15	65.17	55.76	58.18	72.29
SpatialReasoner-R1 fDPO 8B	98.33	98.10	95.28	96.43	91.34	93.64	95.59

	Direct Distance	Horizontal Distance	Vertical Distance	Width	Height	Direction	Quan Acc
Gemini 2.0 Flash [21]	9.45	10.65	26.41	10.52	30.82	54.20	22.43
Llama 4 Maverick [47]	24.48	28.68	34.28	35.71	44.61	58.09	36.72
Gemini 1.5 Pro [63]	14.18	17.21	14.15	19.54	36.09	30.84	21.90
ChatGPT-4o [49]	/	/	/	/	/	60.75	/
SpatialBot-3B [6]	6.00	15.51	8.00	10.52	18.75	39.00	15.62
SpaceThinker Qwen2.5VL-3B [4]	24.32	17.21	59.43	23.27	23.62	32.35	28.97
InternVL2.5-78B [12]	27.70	22.13	41.50	29.32	34.58	62.61	35.25
SpatialReasoner-R1 fDPO 8B	70.95	72.13	74.52	80.45	74.43	94.39	77.30

Spatial Reasoning

