Risultati — 14/10/2025

Third Stage	Dataset	Perception Score	Cognition Score	LLaVA Baseline
Viral	mme	1369.6774 ↓	260.7143 ↓	1510.8/348.2

Third Stage	Dataset	Exact Match	stderr	LLaVA Baseline
Viral	ai2d	0.5214 ↓	0.009	0.548 ↑
Viral	scienceqa_img	0.6837 ↓	0.0104	0.704 ↑
Viral	gqa	0.5709 ↓	0.0044	0.620 ↑
Viral	textvqa_val	0.5204 ↓	0.0068	

Third Stage	Dataset	Accuracy		LLaVA Baseline
Viral	mmmu_val	0.33 ↓		0.353 ↑

Third Stage	Dataset	Accuracy	F1 Score	LLaVA Baseline
Viral	pope	0.8241	0.7931 ↓	0.859 ↑

Third Stage	Dataset	seed_all	image	video	LLaVA Baseline
Viral	seedbench				60.5