

Chess

GPT-4o	0.624	0.644	0.624
GPT-4o-mini	0.610	0.646	0.634
Claude-3.7 Sonnet	0.693	0.656	0.692
Claude-3.5 Sonnet	0.652	0.615	0.551
Claude-3.5 Haiku	0.652	0.550	0.650
Qwen2.5-VL-72B-Instruct	0.542	0.586	0.571
Qwen2.5-VL-7B-Instruct	0.295	0.389	0.386
Qwen2.5-Omni-7B	0.501	0.512	0.448
InternVL3-78B	0.549	0.549	0.560
InternVL3-8B	0.420	0.431	0.454
InternVL-2.5-78B	0.552	0.631	0.635
InternVL-2.5-8B	0.501	0.466	0.450
Llama-3.2-90B-Vision-Instruct	0.495	0.538	0.535
Llama-3.2-11B-Vision-Instruct	0.312	0.446	0.415
gemma-3-27b-it	0.522	0.545	0.517
gemma-3-12b-it	0.454	0.490	0.476

Language

Vision

Vision-Language

Chemistry

0.650	0.573	0.653
0.529	0.409	0.489
0.882	0.816	0.879
0.836	0.813	0.688
0.584	0.534	0.610
0.575	0.559	0.609
0.296	0.443	0.410
0.310	0.307	0.338
0.468	0.446	0.505
0.372	0.289	0.364
0.448	0.404	0.460
0.219	0.328	0.300
0.501	0.363	0.551
0.304	0.294	0.340
0.529	0.499	0.515
0.463	0.446	0.509

Language

Vision

Vision-Language

Music

GPT-4o	0.540	0.328	0.495
GPT-4o-mini	0.348	0.241	0.307
Claude-3.7 Sonnet	0.527	0.396	0.436
Claude-3.5 Sonnet	0.431	0.365	0.344
Claude-3.5 Haiku	0.384	0.307	0.295
Qwen2.5-VL-72B-Instruct	0.425	0.341	0.373
Qwen2.5-VL-7B-Instruct	0.282	0.260	0.280
Qwen2.5-Omni-7B	0.289	0.279	0.306
InternVL3-78B	0.411	0.285	0.361
InternVL3-8B	0.302	0.289	0.278
InternVL-2.5-78B	0.319	0.236	0.283
InternVL-2.5-8B	0.245	0.220	0.221
Llama-3.2-90B-Vision-Instruct	0.331	0.269	0.276
Llama-3.2-11B-Vision-Instruct	0.264	0.245	0.253
gemma-3-27b-it	0.430	0.330	0.354
gemma-3-12b-it	0.395	0.280	0.316

Language

Vision

Vision-Language

Graph

0.725	0.382	0.738
0.735	0.347	0.688
0.876	0.493	0.744
0.739	0.446	0.473
0.559	0.365	0.480
0.647	0.414	0.524
0.340	0.307	0.361
0.351	0.318	0.364
0.674	0.430	0.502
0.434	0.419	0.448
0.471	0.386	0.460
0.331	0.336	0.365
0.409	0.367	0.393
0.275	0.335	0.283
0.581	0.339	0.414
0.521	0.386	0.416

Language

Vision

Vision-Language

