Qwen2.5-VL-72B-Instruct VS gpt-4o-mini Average Score Scatter Plot