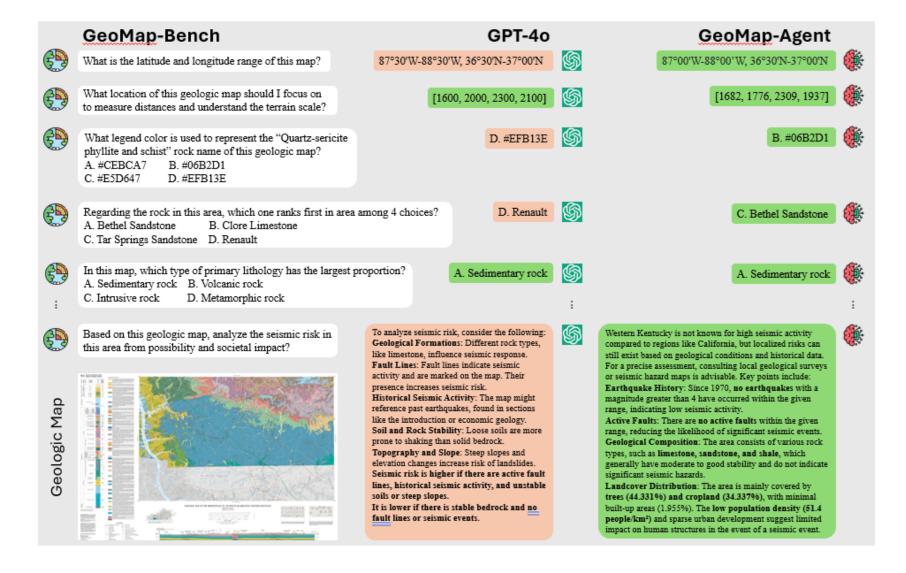# PEACE: Empowering Geologic Map Holistic Understanding with MLLMs

Yangyu Huang, Tianyi Gao, Haoran Xu, Qihao Zhao, Yang Song, Zhipeng Gui, Tengchao Lv, Hao Chen, Lei Cui, Scarlett Li, Furu Wei

Motivation

The Multimodal Language Model (MLLM) is fall short in geologic map understanding due to the challenging nature of cartographic generalization, which involves handling high-resolution map, managing multiple associated components, and requiring domain-specific knowledge.

Contribution

- We developed GeoMap-Bench, the first comprehensive benchmark for evaluating MLLMs' performance in geologic map understanding.
- We introduced GeoMap-Agent, the first AI agent designed for question answering related to geologic maps. It includes modules for information structuring and in-depth analysis using domain-specific knowledge.
- GeoMap-Agent excels existing MLLMs, demonstrating its potential to greatly improve the efficiency and accuracy of geologic map interpretation.
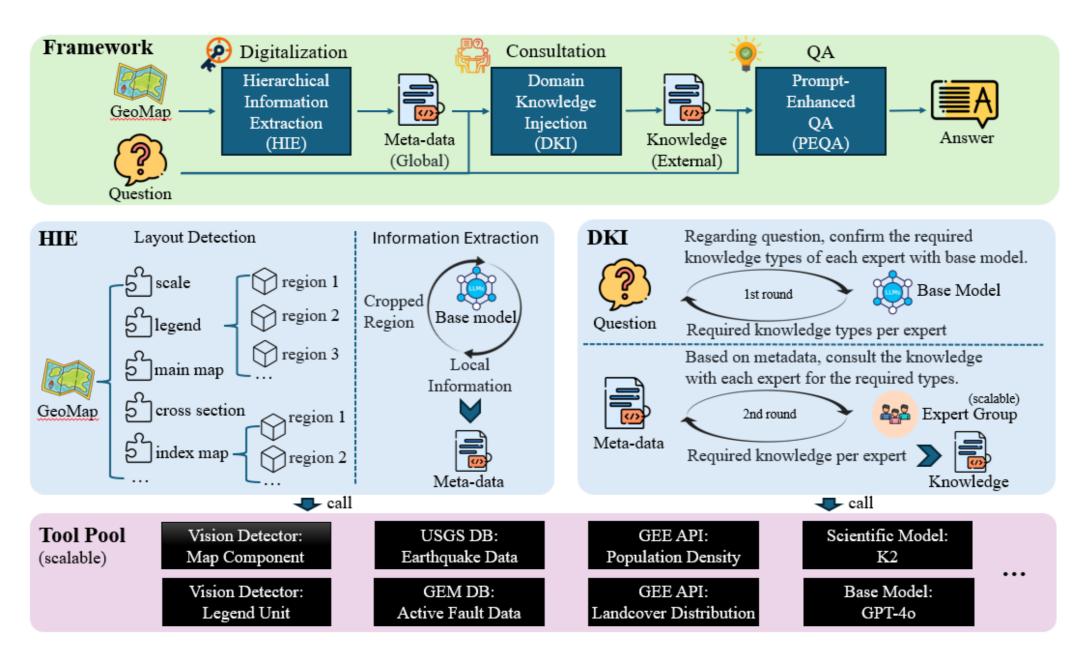
GeoMap-Bench

The GeoMap-Bench is a benchmark designed to evaluate the performance of MLLMs in understanding geologic maps across various abilities, with details provided in below Table and figure.

Property	Description
Source	USGS (English) CGS (Chinese)
Content	Image-question pair with annotated answer
Scale	124 images and 3,864 questions
Resolution	$6,146^2$ pixels on average
Question Type	1. Multiple-choice question 2. Fill-in-the-blank question 3. Essay question
Covering Ability	1. Extracting 2. Grounding 3. Referring 4. Reasoning 5. Analyzing
Defined Task	25 tasks

GeoMap-Agent

The GeoMap-Agent framework includes three modules: Hierarchical Information Extraction (HIE), Domain Knowledge Injection (DKI), and Prompted-enhanced Question Answering (PEQA).

Key Results

We conducted comparison experiments on GeoMap-Bench using various methods, including the publicly available API, like GPT-4o, and open-source models, such as QWen-chat. As shown in below Table, GeoMap-Agent consistently achieves the best performance across different subsets (USGS and CGS) on all aspects of ability, including extracting, grounding, referring, reasoning, and analyzing.

Dataset	Method	Extracting	Grounding	Referring	Reasoning	Analyzing	Overall
USGS Set	Random	0	0	0.250	0.250	0	0.100
	QWen-chat [5]	0.050	0.003	0.253	0.442	0.250	0.199
	GLM-4v-9b [12]	0.050	0.010	0.258	0.212	0.600	0.226
	Idefics-9b-instruct [18]	0.025	0.000	0.247	0.260	0.333	0.173
	Cogvlm2-llama3-chat-19B [15]	0.033	0.000	0.189	0.177	0.067	0.093
	Monkey-chat [19]	0.042	0.010	0.213	0.349	0.267	0.176
	GPT-4o-mini [27]	0.183	0.050	0.278	0.456	0.512	0.295
	GPT-4o [27]	0.208	0.100	0.398	0.494	0.683	0.376
	GeoMap-Agent (Ours)	**0.887**	**0.935**	**0.949**	**0.581**	**0.817**	**0.833**
CGS Set	Random	0	0	0.250	0.250	0	0.100
	QWen-chat [5]	0.000	0.003	0.264	0.334	0.457	0.211
	GLM-4v-9b [12]	0.295	0.076	0.234	0.366	0.468	0.287
	Idefics-9b-instruct [18]	0.000	0.000	0.235	0.134	0.457	0.165
	Cogvlm2-llama3-chat-19B [15]	0.205	0.000	0.236	0.156	0.415	0.202
	Monkey-chat [19]	0.031	0.002	0.248	0.145	0.457	0.176
	GPT-4o-mini [27]	0.204	0.102	0.287	0.474	0.491	0.311
	GPT-4o [27]	0.230	0.157	0.359	0.521	0.542	0.361
	GeoMap-Agent (Ours)	**0.777**	**0.906**	**0.824**	**0.595**	**0.846**	**0.789**
All Sets	Random	0	0	0.250	0.250	0	0.100
	GPT-4o	0.219	0.128	0.378	0.507	0.612	0.369
	GeoMap-Agent	**0.832**	**0.920**	**0.886**	**0.588**	**0.831**	**0.811**

Summary

Geologic maps are crucial for geology, helping with disaster detection, resource exploration, and civil engineering. This study uses Multimodal Language Models (MLLMs) to improve map understanding, aiming to help geologists analyze maps more efficiently. Additionally, we believe the GeoMap-Agent paradigm could be applied to other scenarios facing similar challenges.