

E More Qualitative Results

We present additional visual examples to highlight the geographic reasoning performance. Fig. 13 displays more visual cases involving diverse locations. GRE is capable to generate explainable predictions with robust capabilities in these challenging scenarios. Furthermore, Fig. 14 and Fig. 15 provides comparisons with previous alignment-based methods and existing MLLMs with reasoning capabilities. Our approach exhibits superior image geo-localization results with implicit geographic indicators.



<think>

1. **Architectural patterns:** The image shows a large sculpture resembling a giant pencil surrounded by modern architecture, including a glass-roofed structure that could be part of a train station or transit hub.
2. **Textual clues:** The presence of Italian signage ("Yaris Il Piccolo Genio" and "APITALIA") suggests the location is in Italy. The design of the building and the style of the sculpture are consistent with Milan's urban landscape.

</think>

<answer>

"Latitude": "45.465°", "Longitude": "9.191°"

</answer>



<think>

1. **Textual clues:** Based on the signage of DiBruno Bros., which is known as 'The House of Cheese' and has been established since 1939, it is located in Philadelphia, Pennsylvania.

</think>

<answer>

"Latitude": "39.952°", "Longitude": "-75.165°"

</answer>



<think>

1. **Architectural patterns:** The image shows a street view of a city with tall buildings, including one that is notably the Willis Tower (formerly known as the Sears Tower), which is located in Chicago, Illinois.

</think>

<answer>

"Latitude": "41.882°", "Longitude": "-87.630°"

</answer>

Figure 13: Visual examples of GRE.



Figure 14: Qualitative comparisons with previous alignment-based methods and existing MLLMs with reasoning capabilities. (Lat, Lon) denotes the ground truth coordinates, (\hat{Lat}, \hat{Lon}) denotes the models’ predicted answer, **Indicator** denotes the explicit indicator and **Indicator** denotes the implicit indicator. Notably, GeoCLIP generate five candidates coordinates and select the candidate with the maximum probability score as the answer.

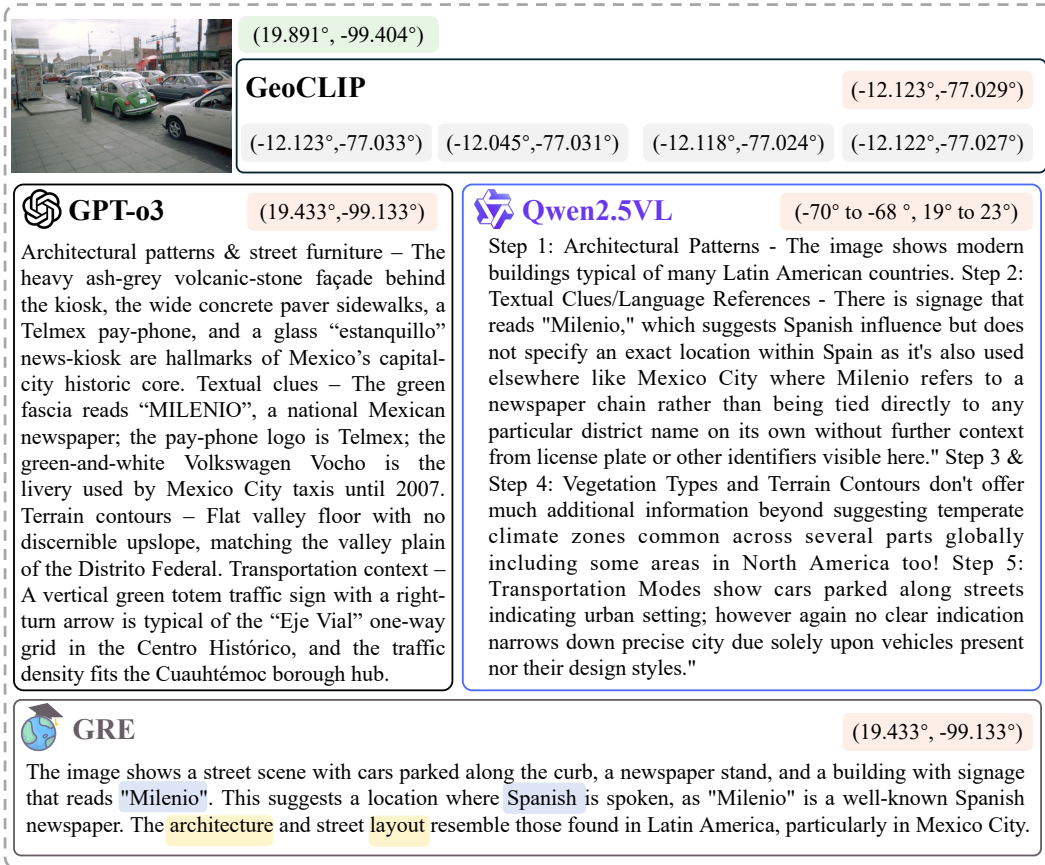


Figure 15: Qualitative comparisons.