This challenge ended up being easier than I had anticipated. I was made aware halfway through the CTF that people were simply using Google Lens to determine where the location of image was. I forgot that Google Lens existed and didn't think to check that while making this challenge lol. Challenge probably should have been rated as easier given that but oh well.

Well played to those who did that, you outsmarted my challenge. Im not gonna let that happen next time...

That being said, the intended solution was much harder and was the reason why it was rated as hard as it was, and it did not involve Google Lens.
The goal was that when given the image, there were only 3 pieces of information that did anything for you.
1. The description of the challenge.
2. The name of the cafe.
3. The partial URL to the left of the image.

You needed to realize that the URL was a .tw Top Level Domain, which means this image was taken in Taiwan.
You then search for ChaCha oolong tea in Taiwan on Google maps, which returns a good number of results.
You needed to then coorelate that with the description of the challenge: the fictional character liked to explore the less touristy sides places. This would mean that mainstream touristy places like Taipei are no longer within the scope of the search leaving only a few places on the outskirts of the island, including Jiufen (the city where the cafe is located).
Then, go into street view and check them all till you find a similar architechtural style to that in the photo and or the only photo sphere that has the right perspective for the given image.
Get the coordinates for that street view photosphere from the URL and there is your flag.
25.1098267,121.8454443