FrancoDetect

Objective

Geoguessr, the well-known game where you can test your geographical knowledge by guessing your location based on Google Street View images, fascinates many people. But what if we take this idea a step further and develop an AI that estimates the geographical position based on real-world images from around the world? That is precisely the goal of this project. We aimed to develop an AI that can guess the location based on images from France[1].

Figure 1: A snapshot from the popular game "Geoguessr," where players attempt to guess the exact location based on Google Street View images.

Background

The game Geoguessr involves the infamous problem of geolocalization. It is the task of determining the coordinates of an image without access to metadata. The reason for the significant difficulty of geolocalization is the sheer range of factors that must be considered. Two images of the same location can differ in terms of time of day, cardinal direction, season, lighting, resolution, camera quality, camera settings, and many other factors.

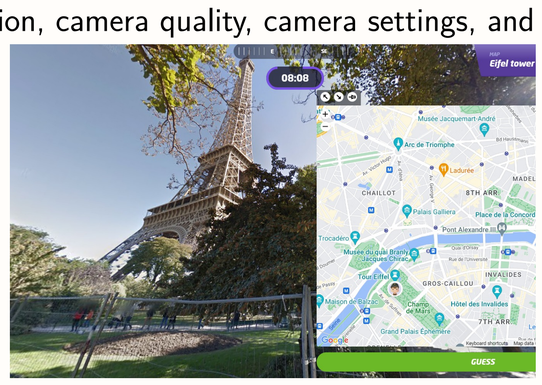

Figure 2: A round of Geoguessr at the Eiffel Tower.

Conversely, images of two different places can have the same vegetation, road signs, architectural style of houses, and still be hundreds of kilometers apart.

Why France?

Figure 3: Google Street View Coverage in Europe

Due to data privacy regulations, road network density, and other factors, not all countries are equally well-covered by Google Street View. France is a country that is almost entirely covered by Google Street View and still possesses a high geographic diversity that an AI can use for location estimation. Therefore, it was an ideal candidate for our project.

Implementation

For the model, we utilize methods of artificial intelligence, specifically, we employ the contrastive learning architecture of neural networks.

Our approach involves first categorizing each image into urban and rural areas, and then separately training models for both cases. Our hypothesis is that urban and rural images require significantly different features (urban architecture vs. rural vegetation). We leverage OpenAI's CLIP [2], to be more precise, a modified version called StreetCLIP [3], to estimate whether the image is urban or rural. Subsequently, we train models on StreetCLIP's image embedding vectors to classify the image into one of the 100 largest cities or one of the 101 départements. For départements, we let StreetCLIP further estimate the arrondissement.

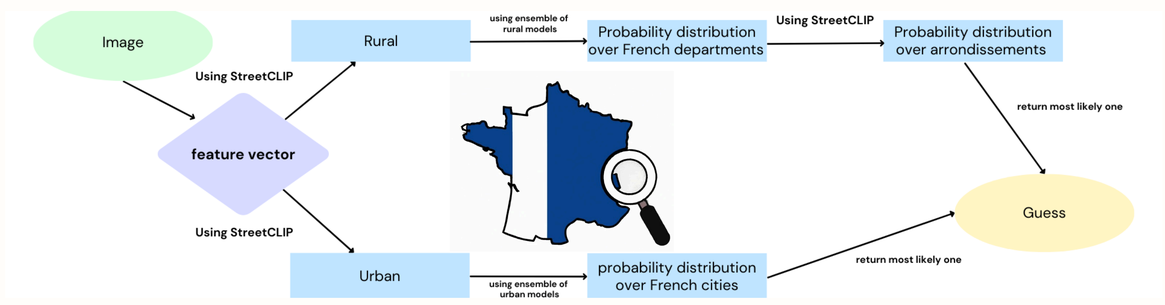

Figure 4: A diagram illustrating our estimation process using StreetCLIP.

Development

We created our dataset for the model by downloading approximately 50,000 images from Google Street View. Accessing the CLIP models was straightforward using Python and Hugging Face. For training the models, we were able to utilize the `scikit-learn` library.

Documentation & Publication

A comprehensive documentation has been provided on GitHub as `README.md`.

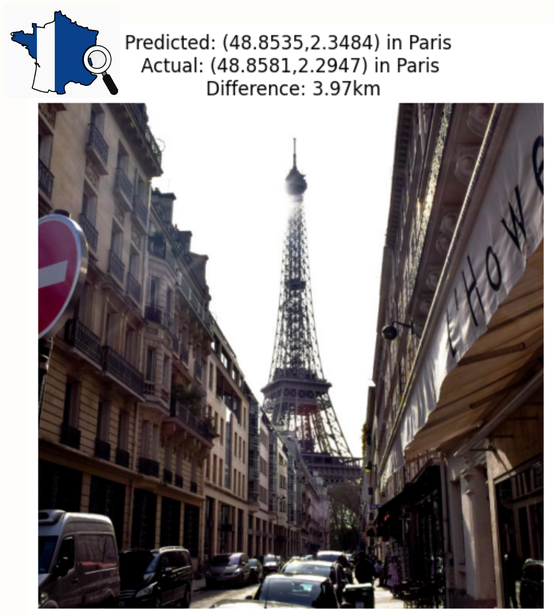

Figure 5: Correct estimation of the Eiffel Tower.

The project has also been published on `GitHub` under the MIT license [4].

Results

The goals outlined in the project proposal have been achieved. We were able to achieve a median accuracy of 75 km and an average accuracy of 115 km. While these results may not be state-of-the-art, achieving such results typically requires significantly more computational resources. Some of our approaches, such as the division into urban and rural areas and feature extraction using the same base model, are original. The MIT license we used allows for both usage and modifications.

Figure 6: Although some images provide few obvious features for humans, FrancoDetect is often able to at least identify the correct region.

Outlook

Geolocalization is a problem that will continue to be better understood and solved in the near future. Significant improvements to the status quo have been achieved by various teams recently [5, 6, 7]. We hope that our project can contribute in a small way to this ongoing progress.

References

[1] `https://github.com/Luka-R-Lukacevic/FrancoDetect`

[2] `https://openai.com/research/clip`

[3] `https://huggingface.co/geolocal/StreetCLIP`

[4] `https://mit-license.org/`

[5] `https://browse.arxiv.org/pdf/2204.13861.pdf`

[6] `https://dl.acm.org/doi/pdf/10.1145/3557918.3565868`

[7] `https://browse.arxiv.org/pdf/2307.05845.pdf`

BND Summer of Code 2023 (Janek Große and Luca R. Lukacevic)