Final Project Proposal

281 Spring 2024 - River Schieberl, Sophia Chen, Sanurak Natnithikarat

Problem Statement

There is a popular online game called GeoGuessr in which the player is dropped into a google street view somewhere in the world and the player has to click somewhere on the map to guess where they are in the world. This is a problem that humans have become fairly good at, often being able to consistently come very close to the true location. Some experts at the game have become so good at the game that they add extra challenges like making the screen black and white, blurry, or only able to see part of the screen.

Humans employ a variety of techniques that often incorporate knowledge of the “real world” such as vegetation, building architecture, and road configuration. Our project’s goal is to train models to play this same game, but we will simplify the task to match our available data.

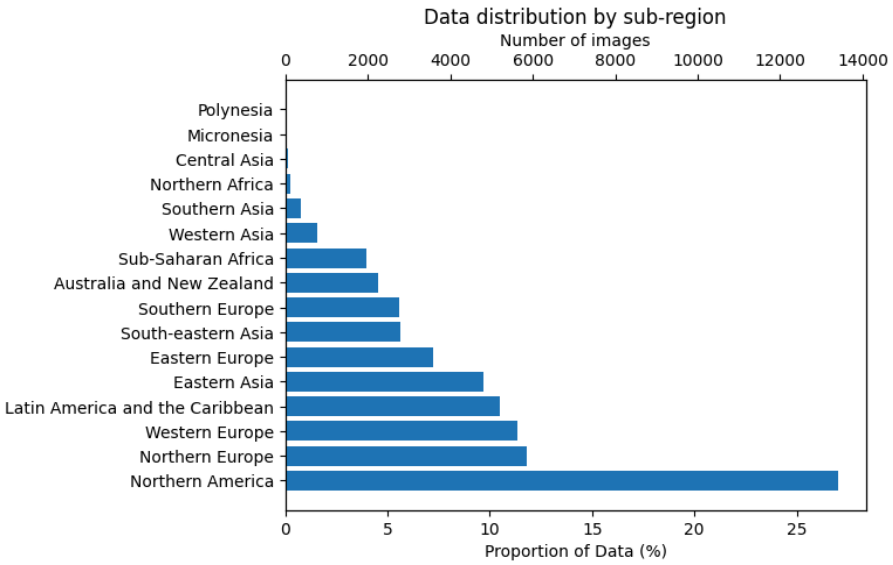
Data Source

https://www.kaggle.com/datasets/ubitquitin/geolocation-geoguessr-images-50k

Data Label Description

The dataset linked above contains around 50,000 images from 129 countries. We felt that there would not be substantial variation between countries to warrant having so many classes to predict, so we decided to merge countries together into subregions to cluster similar regions.

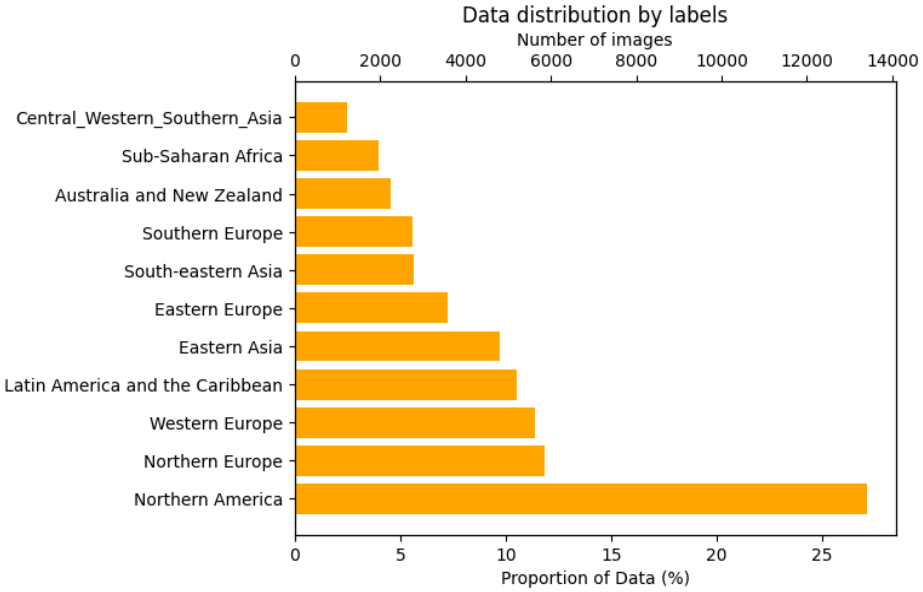
A map showing the subregions that the countries are grouped into.
This map is very similar to the actual groupings, but there are some minor differences.



A histogram showing the number of images for each subregion.

However, we have identified a strong data imbalance persisting after checking the dataset distribution based on these labels. Northern America contains nearly a quarter of the data and Polynesia contains a single digit number of images.

We propose dropping Polynesia, Micronesia, and Northern Africa and grouping Central, Southern, and Western Asia. The more balanced distribution of labels is shown below, with the smallest class containing 1,224 images. We are exploring downsampling, augmentation, and weighting strategies to further balance the dataset.



A histogram showing the number of images for each subregion after modifying smaller classes.

Example Images

An example image from each subregion.

Image Description
The image size seems to be similar across all classes (1536, 662) pixels on average with some slight variations. With all images containing the same artifacts from the game overlay on the edge and the center, we plan to crop all of the images into a consistent size while eliminating as much as the game overlay interface as possible.

An example image showcasing the game interface artifacts to hide.

Relevant Features
One of the most relevant features that we expect to be useful for classification is the color of the image. Regions that are drier will tend to have more browns compared to regions that are less dry which will have more greens.

We expect information about buildings to also be a useful feature. Regions with more buildings will tend to be in denser areas like Europe while regions with less buildings will be in sparser areas like Northern America or Australia. The edges and shapes of the buildings will also be useful information as architecture varies significantly between regions.

The texture of the image will also likely be a useful feature as we expect different regions to have very different textures relating to the climate and type of cities.

One feature that we're interested in exploring is measuring the lumens in the picture. We hypothesize that brighter regions will be more likely to get sunlight so there could be some correlation.

The histogram of oriented gradients is another feature that we are looking forward to researching as it has potential for describing the structure and shape of an object. Since we anticipate different regions to have sufficient variability in their structures, this could lead to succinctly synthesizing the information in the images.