

Southeast Asia GeoGuessr: Visual Country Identification

Mariah N Cornelio

Abstract

This project explores whether a deep learning model can distinguish between visually similar Southeast Asian countries, specifically the Philippines, Indonesia, Malaysia, and Thailand, using street-level imagery. By creating a custom dataset and applying transfer learning with CNNs, the goal is to build a classifier that can accurately identify the country shown in each image.

Introduction

GeoGuessr is an online geography game that challenges players to identify locations based on street-level imagery. As someone who enjoys playing GeoGuessr to learn about different cultures and geographies, I've noticed that while identifying countries is often manageable, distinguishing between visually similar Southeast Asian nations is particularly difficult.

Motivation



One of these images depicts El Nido, Philippines, and the other depicts Phuket, Thailand. Can you guess which is which? These countries share common visual traits like tropical vegetation, urban density, and similar infrastructure, making it hard to tell them apart without clear markers like flags or linguistic text. This project explores whether deep learning can detect subtle visual cues that distinguish these countries. The model will learn to classify street-level images by country, using a custom-built dataset sourced from Google Maps API and Google Maps Street View.

Problem Formulation

While some global models exist for geographic image prediction (like GeoSpy and AI-based world location predictors), few focus on fine-grained classification within a specific region like Southeast Asia. This project targets that gap by training a model to classify images from four visually similar countries: the Philippines, Indonesia, Malaysia, and Thailand. Since no suitable dataset exists, I will create one using manual screenshots and the Google Street View API, collecting 200–300 images per country. The task is an image classification problem using convolutional neural networks (CNNs) with transfer learning from architectures like MobileNetV2 or ResNet50. The dataset will be split into training, validation, and testing sets (70/15/15), with image augmentation applied during training. Model performance will be evaluated using accuracy, precision, recall, F1-score, and a confusion matrix for each country. ***The goal of this project is to build a model that takes an image of one of the listed countries as input and outputs the country name along with the respective metrics.***

Workplan

The project will begin with collecting 200–300 images per country using screenshots from Google Maps and the Google Street View API. Images will be organized into folders by country and resized for consistency. The dataset will then be split into training, validation, and testing sets, and data augmentation will be applied during training to improve model generalization. I will implement a transfer learning approach using a pretrained CNN model such as MobileNetV2 or ResNet50. After training, I will evaluate the model using accuracy, precision, recall, F1-score, and confusion matrices. The final deliverable will be a notebook containing the full pipeline, along with test predictions and possibly some Grad-CAM visualizations to interpret the model's decisions if time permits.

Collect & organize images → Preprocess data → Split dataset → Build & train models → Model evaluation & Grad-CAM → Improve if needed