# CSCI5922 Neural Networks Group Project: GeoguessrLSTM Project Proposal

**Nirvan S P Theethira**
nith5605@colorado.edu

**Dheeraj Ravindranath**
dhmu3474@colorado.edu

## Abstract

Geolocation is the identification or estimation of real-world geographic location using location based data. The location based data used for geolocation includes but is not limited to satellite images, sensor data, traffic movement data, pollution data etc. While humans cannot effectively guess locations based on sensor data, well travelled humans are good at guessing the location of where the image was taken. This project hopes to use neural networks to emulate the results of a human when it comes to geolocation of google street view images.

## Motivation

The motivation for this project comes form the game called Geoguessr [1]. The game presents players with street view images (see Figure 1), which the player, using any pre-existing knowledge, has to guess the location of the image. The game scores the user based on how closed the latitude and longitude of the guess was to the actual latitude and longitude of the location of the street view image. This is an extremely challenging task for a machine that has no prior knowledge that an avid traveller possess, therfore making this an exciting problem to tackle using neural networks. While previous projects have aimed to geolocate prominent landmarks of locations such as the Eiffle tower, The Taj Mahal etc. this project aims to use generic street view images. As seen in Figure 1, there are mountains and red stone buildings which can hopefully be used by a deep neural network to ascertain the image is from boulder.

Figure 1: Google street view of pearl street.

Preprint. Under review.

## Data collection

The first step involves defining the region for data collection. For this project the map of mainland United States will be chosen as the landmass has diverse geographical features and has a plethora of street view images available. The US mainland boundary shape file as of 2018 is published by the unites states census bureau [2]. The next step is to split the map into logical grids as shown in Figure 2. Data points will be scrapped from multiple locations in each grid. As street view offers a 360 degree view per location, multiple images will be collected at each location. The google street view static API [3] will be used to collect image data.

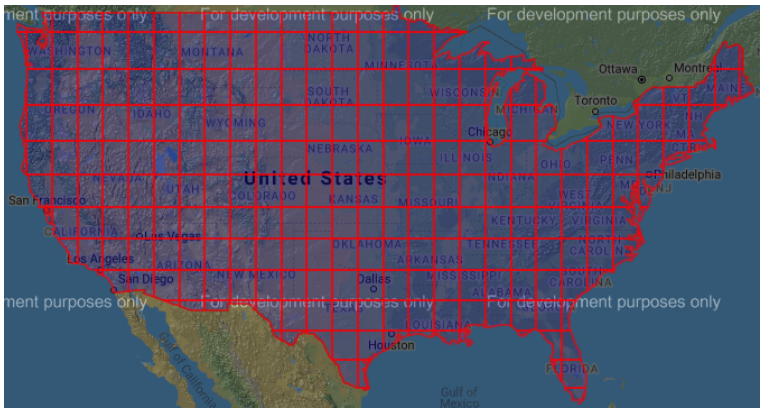


Figure 2: US mainland map split into grids. Data will be collected at multiple random locations in each grid.

## Proposed Neural Network

The neural network structure consists of a CNN/LSTM combination (see Figure 3). This idea was inspired by PlaNet paper [4]. Here a single training epoch will consist of multiple images obtained from a single 360 degree sweep of a location. To process multiple images in one training epoch, the list of images are fed into CNN's that give the vector representations of the list of images. These vectors are then fed into a sequential LSTM and the output is soft maxed across grids from which data is collected. Google Colab GPU's will most likely be used for training.

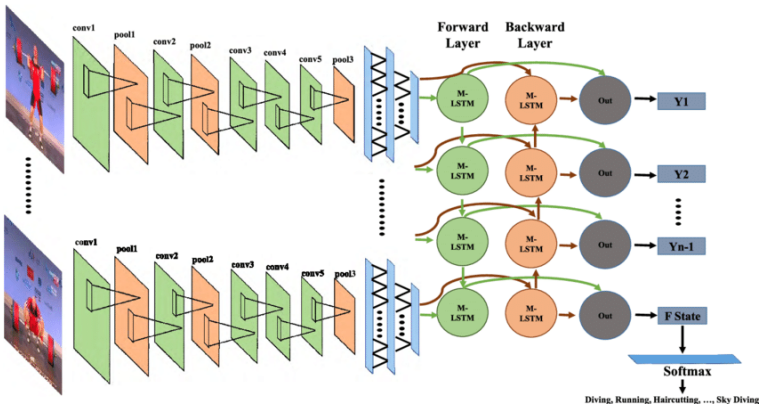


Figure 3: Structure of the proposed deep learning model.

## Evaluation

The evaluation for the project will be done based on the euclidean distance between the predicted location grid and the actual location grid. The test set will be processed using the above mentioned machine learning techniques to get the softmaxed output across grids. The top N softmaxed outputs are considered. Here N will be a hyper parameter. Out of the top N, the one that gives the grid with the least euclidean distance from the target location will be chosen. This way the models can be evaluated with the euclidean distance across the test set being summed up. The model with the least sum will be considered the best. Human player scores will be used as baselines to compare with model scores.

## References


[1] Anton Wallén. Web-based geographic discovery game. `https://www.geoguessr.com/`, May 2013.

[2] Unites States Census Bureau. Cartographic boundary files - shapefile. `https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html`, 2018.

[3] Google Maps Platform. Street view static api. `https://developers.google.com/maps/documentation/streetview/intro`, 2020.

[4] Tobias Weyand, Ilya Kostrikov, and James Philbin. Planet - photo geolocation with convolutional neural networks. *CoRR*, abs/1602.05314, 2016.