EUGuess - EU Countries Geolocalization given Random Images with Several Perspectives

Carlos Pérez, Luis Pérez and Jorge Garcelán

[jgarcelan3, lperez63, cjimenez41]@gatech.edu

INTRODUCTION

Nowadays, image geolocalization is still a challenging task in the field of computer vision. Our proposal is to tackle image geolocalization as a classification task using convolutional neural networks, classifying random localizations generated from random coordinates with their given country, within 32 European countries. Furthermore, we also make a comparison between training our networks with 1 or 3 perspectives of the same locations, getting slight differences, especially in test accuracy.

EUGuess confronts several handicaps, but the most significant is the size of the imagery. The complexity of the image generation and the lack of high performing computational resources leads to a small dataset of 14000 locations with 42000 images taking into account the 3 perspectives. Despite the inconveniences, EUGuess significantly improves the accuracy of guessing the location's country randomly.

Figure 1. Random Image from Albania

APPROACH

After generating the dataset, we have decided to resize all the images to 64x64 pixels. The main reason for carrying out this action is the huge size of the images from Google Street API (640x640 pixels), that lead to problems related to space of RAM and execution time.

The next step we made was to save a copy of the normalized tensor of every image, stored as a pickle object, for optimizing the process of loading data of images.

After the processing of images, tensors and perspectives we have tested a considerable number of models from the torchvision library. EUGuess has focused its training in four main models: Resnet50, EfficientNetB4, GoogleNet and VGG16. Additionally, we have tested our own CNN including several Convolution layers with ReLU as activation function, MaxPooling and BatchNormalization between layers and finally fully connected layers with dropout.

Figure 2. Random Image from Albania

EUGuess tested the models mentioned above considering one or several perspectives. In order to carry out this test, we have taken only 14258 images as the dataset has that exact number of locations and by doing this models will train with same number of images.

Our final step was to test the models we have considered for the project with the whole dataset of images. We have tuned hyperparameters, made use of different optimizers such as Adam or SGD, *CosineAnnealingWarmRestarts* scheduler and used categorical entropy loss as the loss function of EUGuess.

DATA COLLECTION

It is commonly known that 70% of the time of a data scientist's work is spent on data. Preparing, processing, loading and cleaning the data are some of the most important tasks in machine learning. In this project, a great part of the work is related to the creation and implementation of proper datasets with locations along Europe, generating random images from specific EU countries. using the API of "Google Street View".

In order to generate an image we inserted random coordinates of longitude and latitude as inputs. If that coordinates lead to a valid location of Google Street View, it generates a jpg 640x640 pixels image. Moreover, to generate the three desired perspectives we inputted 3 different angles of the camera: 0º, 120º and 240º (Figure 5).

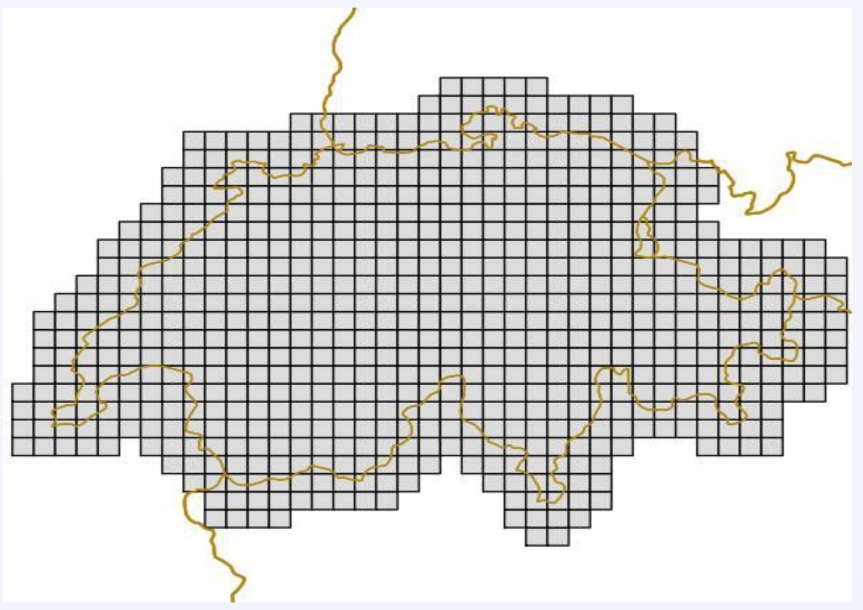

Figure 3. Switzerland shapefile

In order to generate random coordinates within the range of areas of the different countries, we decided to use shapefiles (*Figure 3*), which are geospatial vector data for geographic information system (GIS) software. Shapefiles are geolocated grids of such a country, therefore we needed to establish maximum and minimum bounds for each country, so we can take random numbers for latitude and longitude, and thus, generate a random location within that bounds.

We have decided that the number of images per country will vary depending on population, economy, how manageable is for Google Street View to generate a new location. However, we are specially taking into account the area of the country. In this way, we end up with unbalanced classes, this is because a certain number of images does not represent equally two countries with big differences in surface.

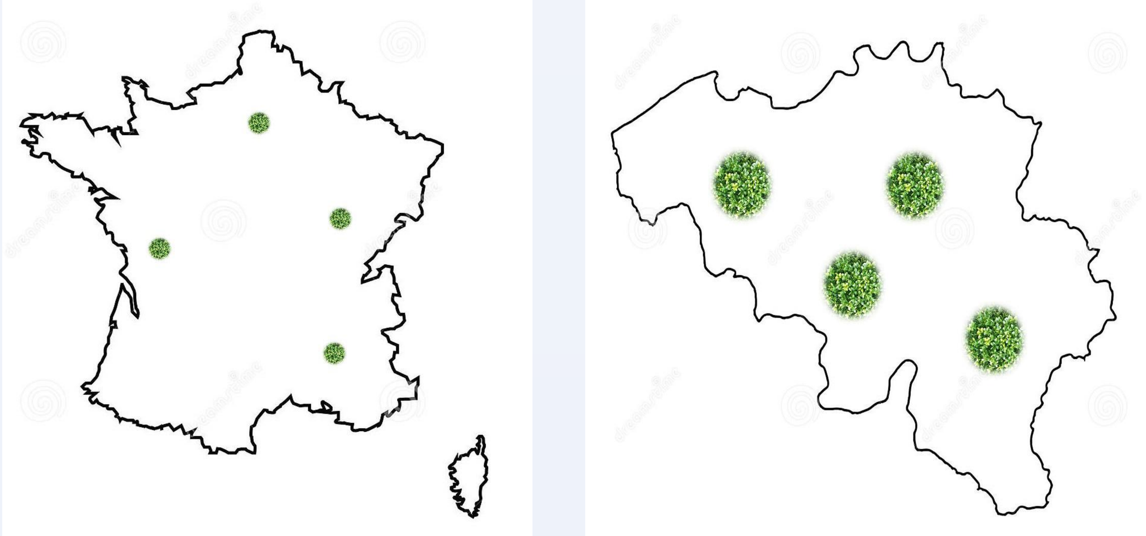

Figure 4. Points representation in countries shape

For instance, France's area is 248,573 m2 whereas Belgium's surface is 11,849 m2, meaning that 4 random images are much less representative for France than for Belgium as shown in Figure 4 where the location points are exaggerated.

Therefore, we decided to classify the countries into three categories: small, medium and large. For the first category we have generated 123 locations (369 images) per country, for the second category 369 locations (1107 images) and for the last category 860 locations (2580 images).

Country	# Locations	# Images
Spain	860	2580
Germany	860	2580
Ireland	369	1107
⋮	⋮	⋮
Croatia	369	1107
Slovenia	123	369
Luxembourg	123	369

Table1. Locations and Images distributions

RESULTS

Our first important optimization results was to transform jpg files into tensors and saving them as pickle format. This step allowed as to load the data of images needed significantly faster. The load of the whole dataset of jpg files took between 2 and 3 hours while loading tensors took less than 5 minutes.

Figure 5. Three perspectives from Albania location

The main experiment of our project was to test different Deep Learning models with the whole dataset of images including three perspectives. The four models we have considered due to the accuracies obtained when testing around 10 different models are: GoogleNet, EfficientNetB4, VGG16 and Resnet50.

After deciding which models we were going to train we followed an exhaustive search of different hyperparameters, optimizers and schedulers in order to improve EUGuess accuracy. In the following two figures different models outputs are represented:

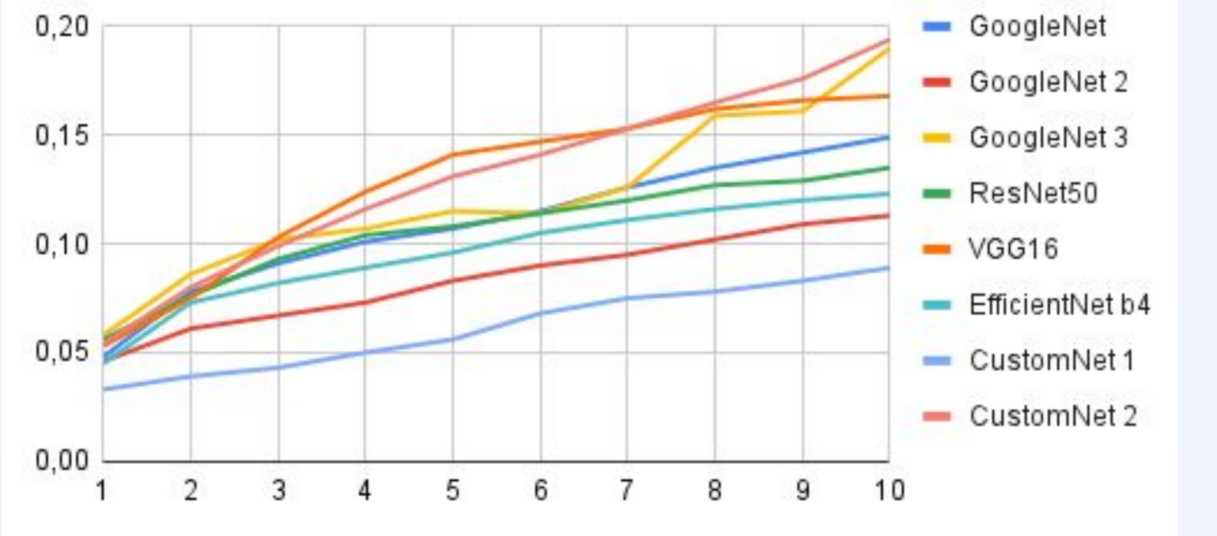

Figure 6. EUGuess training accuracies

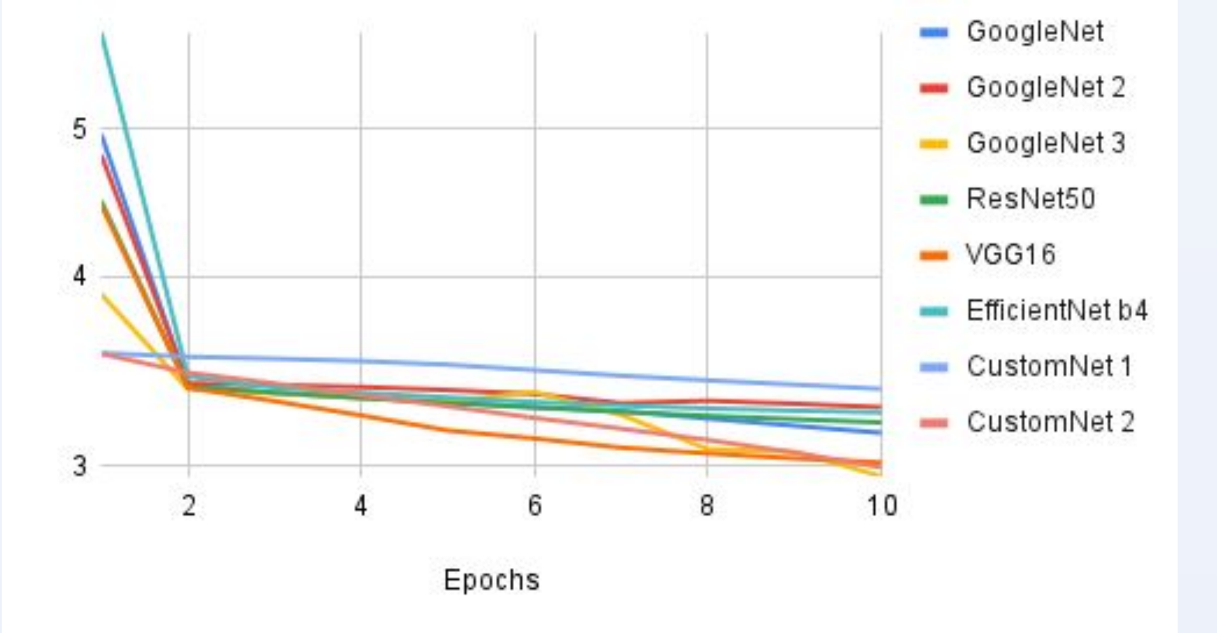

Figure 7. EUGuess training log losses

The main information extracted from these plots is the better performance of GoogleNet when hyperparameters are well tuned as it is the case of GoogleNet3. Due to these results, we decided to explore deep into this model looking for a higher accuracy. In order to achieve this goal, we tested several schedulers with different parameters and tuning the hyperparameters models. Finally, using optimizer SGD with *CosineAnnealingWarmRestarts* EUGuess has reached a 25% test accuracy with the optimal hyperparameters.

The final experiment was to test EUGuess with completely new images, to do so, we have generated 1000 more images of the 32 countries. In this way, none of the pictures is a perspective of the training imagery. So we avoid the natural overfitting of our proposal. From a 25% accuracy testing with perspectives, we have gotten 16% with the renewal test set.

The previous test accuracy is almost as exact as EUGuess performance when it considers a unique perspective. With these complementary test we checked that when considering three perspectives, the dataset is significantly larger and the accuracies obtained were around 50% higher.

CONCLUSIONS

In the barplot of Figure 8, we can see the performance while training the different models using 3 perspectives. One can conclude that the best CNN is GoogleNet 3. Getting 20.9% as you can see in *Figure 8*. After hyper-parameter tuning we reached 25.1% as you can see in *Table 2*.

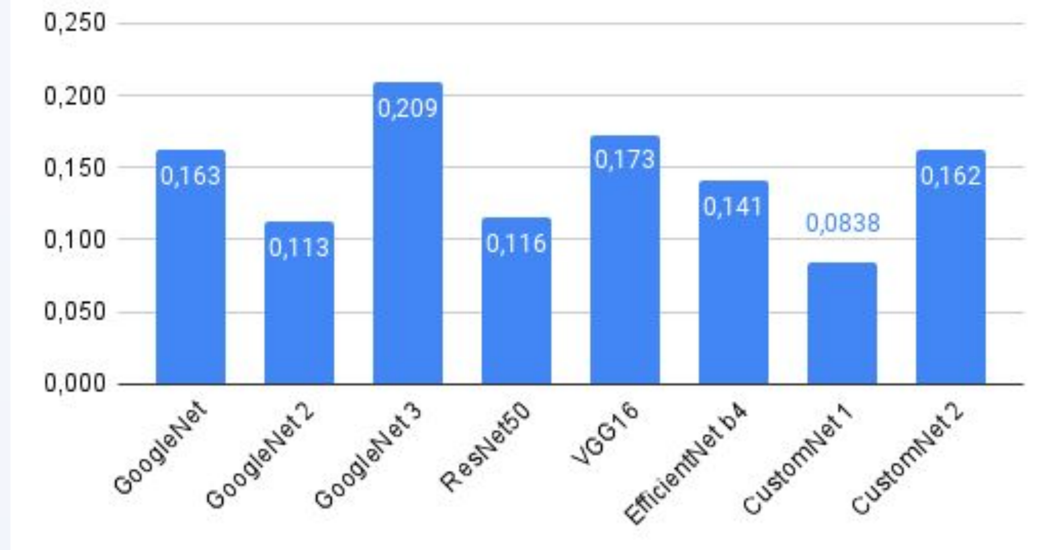

Figure 8. EUGuess test accuracies

Our presented work proves that completely random locations from Google Street View can serve to build a model able to guess the country of 1/4 of these images, just being trained with no more than 34k photos, 11,2k locations in total. This high accuracy comes from the fact that in the test set we can find some of the perspectives that have obviously similar features to the other perspectives of the same location used when training.

With our own neural network (CustomNet) we got more than 16% of accuracy which is a very decent result too. Something we want to highlight as well is the performance of EUGuess with just 1 perspective, achieving an accuracy close to 15% being trained with just 11,2k images of different locations, a very small imagery.

Stating this, EUGuess performance in a real test with images that does not belong to the initial 42k would be of at least 16% as we proved in the experiments. In *Table 2*, there are shown EUGuess methods and accuracies and related previous works: *Im2GPS* and *PlaNet*. These frameworks use different approaches but the objective is similar. In parenthesis in Table 2, we can observe how many images where use to train the models. It should be noted that these proposals are world-scale not only Europe.

Method	Country Accuracy
Im2GPS (orig)	23 %
Im2GPS (new)	35.4 %
PlaNet (900k)	21.6 %
PlaNet (6.2M)	45.6 %
PlaNet (91M)	53.6 %
EUGuess - GoogleNet (42k)	25.1 %
EUGuess - CustomNet (42k)	16.2 %

Table 2. Related Works´ accuracies comparison

REFERENCES

- GEOGUESSR
- REVISITING IM2GPS IN THE DEEP LEARNING ERA., *Vo, N., Jacobs, N., Hays, J*
- PLANET-PHOTO GEOLOCATION WITH CONVOLUTIONAL NEURAL NETWORKS., *Weyand, T., Kostrikov, I., Philbin, J*
- EEA, EUROPEAN ENVIRONMENTAL AGENCY
- MAP PROJECTIONS FOR EUROPE, EUROPEAN COMMISSION JOINT RESEARCH CENTRE.
- LARGE SCALE VISUAL GEO-LOCALIZATION OF IMAGES IN MOUNTAINOUS TERRAIN., Baatz, G., Saurer, O., K¨oser, K., Pollefeys, M
- CNN ARCHITECTURE FOR WEAKLY SUPERVISED PLACE RECOGNITION., *Arandjelovic, R., Gronat, P., Torii, A., Pajdla, T., Sivic, J*
- LANDMARK CLASSIFICATION IN LARGE-SCALE IMAGE COLLECTIONS., *Li, Y., Crandall, D.J., Huttenlocher, D.P*
- LOCALIZING AND ORIENTING STREET VIEWS USING OVERHEAD IMAGERY., *Vo, N.N., Hays, J.*
- ACCURATE IMAGE LOCALIZATION BASED ON GOOGLE MAPS STREET VIEW., *Zamir, A.R., Shah, M.*
- LARGE SCALE VISUAL GEO-LOCALIZATION OF IMAGES IN MOUNTAINOUS TERRAIN., *Baatz, G., Saurer, O., K¨oser, K., Pollefeys, M*

CONTACT

We are three Spanish exchange students from *Universidad Carlos III of Madrid.*

Jorge Garcelán Gómez - jgarcelan3@gatech.edu

Luis Pérez Olid - lperez63@gatech.edu

Carlos Pérez Jiménez - cjimenez41@gatech.edu

*With the QR code you can access an interactive game of the project. The purpose is to show what kind of images EUGuess has operated with.

SCAN ME