Using Machine Learning Methods to Identify and Classify the Regions and Projections of Online Maps

Jialin Li* and Ningchuan Xiao

Department of Geography, Ohio State University
*Email: li.7957@osu.edu

Abstract

In this paper, machine learning methods are developed to identify and classify maps from online sources. Given any image, our methods can be used to answer three questions: is it a map, which geographic region is this map about, and what projection is used in the map? A total of 1000 images are collected using Google Images to provide data for non-maps and maps that cover different geographic regions with different projections. Three machine learning methods are tested: multilayer perceptrons (MLP), support vector machine (SVM), and convolutional neural network (CNN). Each method is evaluated by how accurate it can be used to correctly identify and classify a given image. For map identification, all three methods can be used to obtain results with high accuracy, while for map classification based on region and projection, only SVM and CNN perform well.

Keywords: Cartography, Machine Learning, Map Understanding, Image Classification.

1. Introduction

The overall premise of this research is that artificial intelligence will continue to evolve so that computers will be able to understand human maps (and without a doubt other artifacts) and help us make better maps. This paper represents a preliminary step toward this direction with an aim of testing the capability of machine learning methods in recognizing online map images. We set three specific objectives. First, we develop methods that can be used to tell if an image is a map or not (identification). A successfully identified online map can then be further processed using machine learning methods, and our next two objectives are related to the classification of the map: to tell which geographic region the map image is about and to tell which projection is used in the map.

To the best of our knowledge, there has been few research about automatic map identification and classification in the cartography or computer science literature. In the paper, we extended machine learning methods that are used in other fields to our research. In the remainder of this paper, Section 2 discusses the related methods in machine learning, section 3 describes the data used and experiment results. We conclude our research in Section 4 with a discussion of experiment results and their implications.

2. Data and Methods

A total of 1000 images of maps and non-maps are collected using Google Images. Specifically, 200 non-map images of different themes are collected (Figure 1). For map images, we collect 400 maps for four geographic regions (the world, China, South Korea, and the United States), 100 for each region (Figure 2), and 400 world maps for four different projections (Equirectangular, Mercator, Miller, and Robinson projection), 100 for each projection (Figure 3). All images are resized to a fixed size of 120 by 100 pixels using nearest neighbor resampling.

A huge body of literature on automatic image identification and classification has been established in fields such as facial recognition (Lemley et al., 2016) and remote sensing image processing (Bentes, Velotto, and Lehner, 2015; Qian et al., 2015). Methods developed in these fields can be used to address the questions raised in this paper. More specifically, we identify and test three methods that have potential in this research: support vector machine (SVM), convolutional neural network (CNN), and multilayer perceptron (MLP).

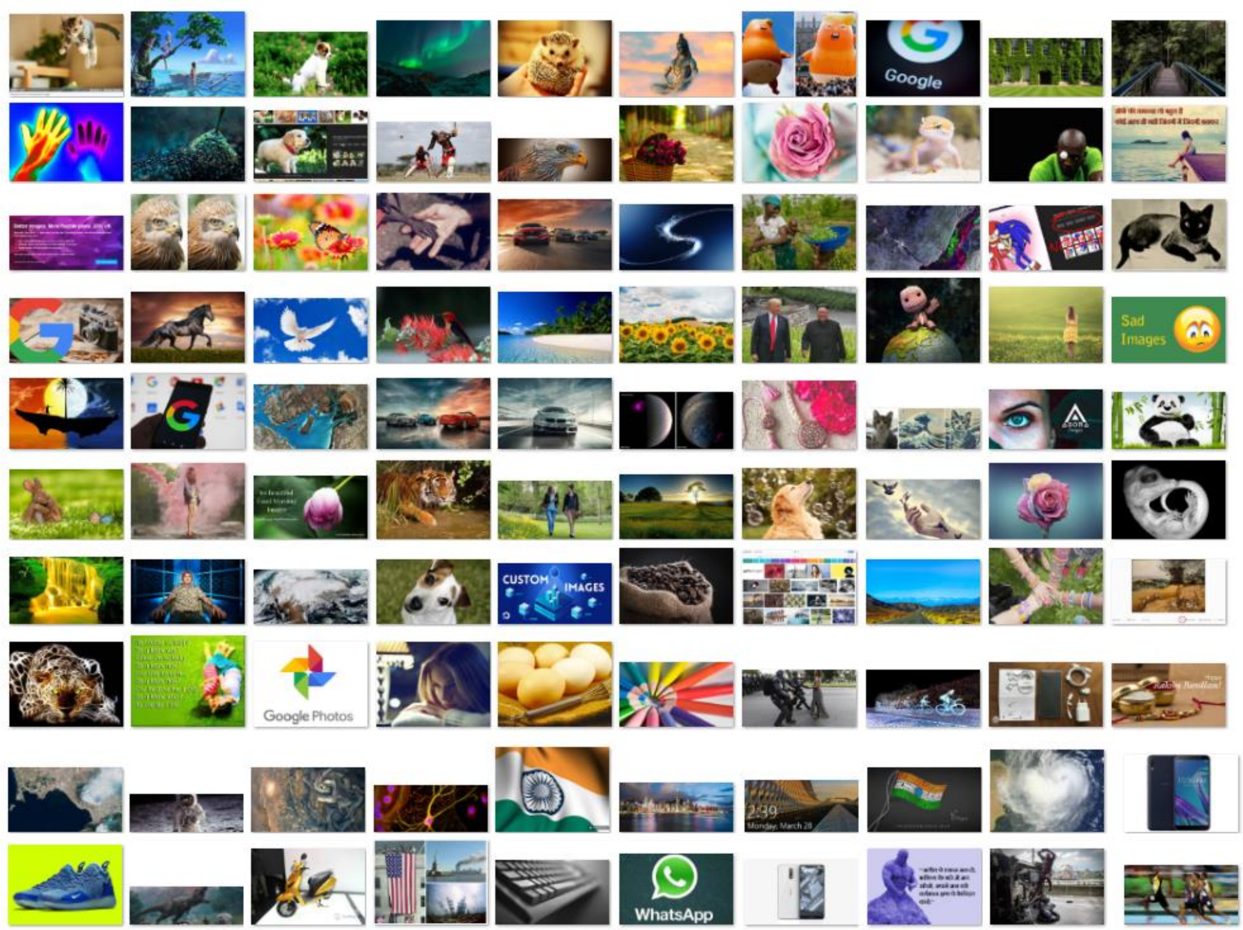

Figure 1: Examples of non-map images

In an SVM, the inputs are represented as points in a multidimensional space, and we find an optimal hyperplane that can be used to separate the points into classes. A hyperplane is optimal if it exhibits a maximum margin between the classes and minimal misclassification error (Cortes and Vapnik, 1995). In our research, each image is a point in a space of 3x120x100 dimensions (where 3 is the three bands of red, green, and blue). When the points are linearly inseparable, kernel functions are used to transform the original data into a higher dimensional space so that the optimal hyperplane (linear) can be found (Haykin, 2009). Various kernel functions have been developed to perform such transformation (Duda, Hart, and Stork, 2001). Four commonly used kernel functions (linear, polynomial, radial basis function, and sigmoid) are applied in our research. The kernels and SVM in general reply on multiple parameters to configure and we use 5-fold cross validation (Haykin, 2009) to find the combination of the parameters for the best performance. For the three methods, two classes (map and non-map) are used for the identification tasks, whereas for the classification tasks we have 4 classes for the four regions, and 4 classes for the four projections.

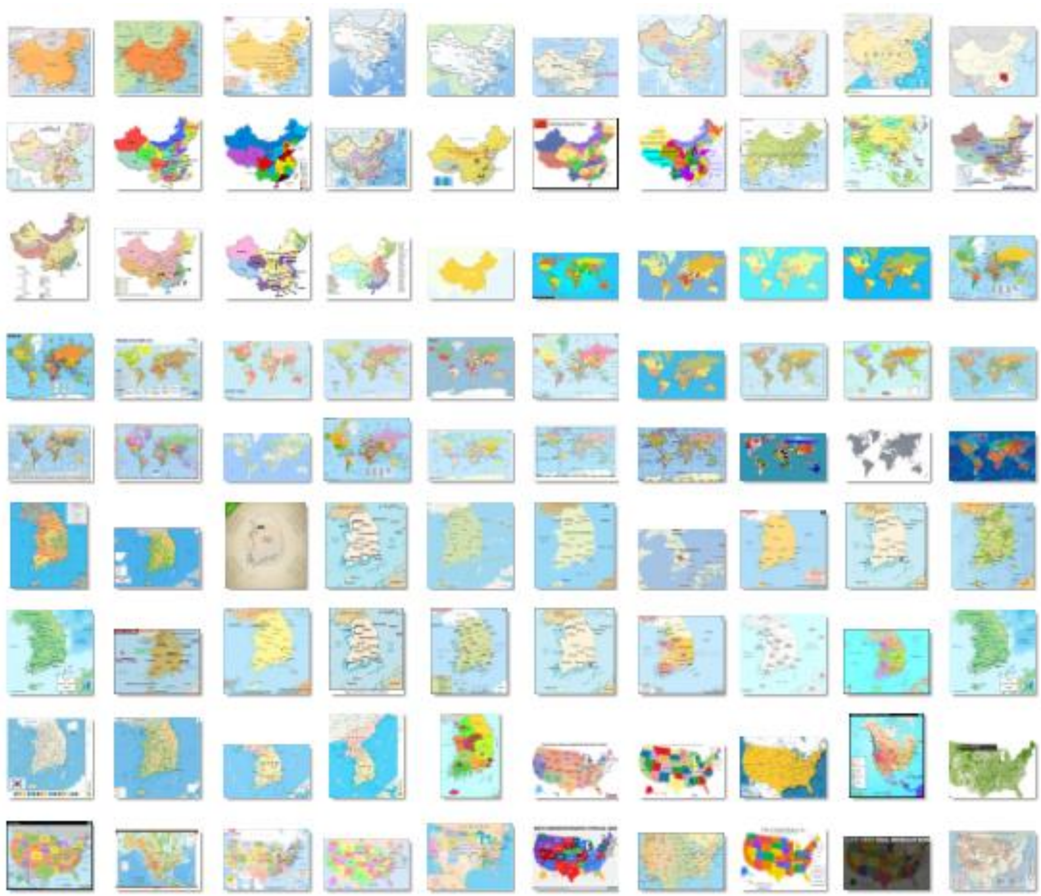

Figure 2: Examples of maps for four regions

MLPs and CNNs are examples of artificial neural networks (ANN). An ANN is a network of layers of nodes, where each node is called a perceptron that transforms a set of inputs into an output. The inputs are summed using a set of weights and the perceptron is activated to generate an output value of 1 if the sum exceeds a certain value. An ANN has a number of layers of nodes, where two of the layers are the inputs and outputs from the data, and other layers in between the input and output are called hidden layers. The goal of developing an ANN is to find the optimal set of weights that minimize the overall error between the ANN output and the data. Different methods have been proposed to find the weights (Haykin, 2009). MLPs are a type of classical neural network that uses backpropagation algorithms to find the optimal weights. While MLPs are effective in many applications (Zanaty, 2012; Sibanda and Pretorius, 2011), they have a relatively small number of layers of node (shallow networks), which makes it hard for them in image processing. CNNs address the shortcomings with MLPs because CNNs use a feed-forward architecture that allows more layers of nodes (deep networks) to be (LeCun, Bengio, and Hinton, 2015). More specifically, a CNN has convolutional layers which are designed to discover spatial structures in the data (Goodfellow, Bengio, and Courville, 2016). In our research, MLPs are used as a benchmark and are compared with CNNs.

Figure 3: Examples of maps in four projections

The inputs in both MLPs and CNNs are also the pixel values in three bands of each image: there are 3x120x100 input nodes. In order to obtain the optimal parameter setting of the number of hidden layers and nodes, we try the number of hidden layers from 1 and increase it to see the trend of outcoming testing accuracy. Similarly, for a fixed number of hidden layers, we increase the number of nodes in each hidden layer from a small number until the accuracy begins decreasing. Experiments are also conducted to test the influence of training size on the performance of the three methods for map identification and classification. Additionally, training time for all sets of experiments is recorded and compared. Parameters in MLPs and CNNs are shown in Table 1 and 2 respectively. A Python package, Keras, running on top of Tensorflow, is used to accomplish MLPs and CNNs, and for SVMs, a Python library called LIBSVM (Chang and Lin, 2011) is applied.

Parameters	Values or Choices
Activation function	Rectified linear unit
Optimizer	Stochastic gradient descent
Learning rate	0.01
Number of epochs	100

Table 1: Parameters in MLPs

Parameters	Values or Choices
Activation function	Rectified linear unit
Kernel size	5*5
Pool Size	2*2
Strides	2*2
Optimizer	Stochastic gradient descent
Learning rate	0.01
Number of epochs	100

Table 2: Parameters in CNNs

3. Experiments and Results

For the map identification task, we use a total of 400 images, including 200 non-map images and 200 map images that consist of 50 images randomly selected from each of the four regions. The 400 regional maps are used for the region classification task, and the 400 projection maps for the projection classification task. For all the tasks, 90 percent of the images are used for the training purposes and 10 percent are used as testing data.

For map identification, the experiment results of the three methods suggest that all three methods can be used to obtain high accuracy rates (Table 3). SVMs and CNNs can correctly identify testing images with a 100 percent accuracy rate. For region classification, the accuracy rate of classification by MLPs is low, while SVMs and CNNs can be used to classify maps correctly with a 100 percent accuracy rate (Table 1). Similarly (and as expected), the results of MLPs are of low accuracy for projection classification, while SVMs and CNNs can still classify maps with relatively good results (Table 1). It can be noted that projection classification appears to be more challenging than map identification and region classification since none of the methods can classify testing images with an accuracy rate higher than 90 percent.

In general, polynomial kernel function is the best kernel for SVMs in the three tasks. A CNN with two convolutional layers (containing 32 and 64 nodes respectively) generates best results in map identification and region classification, and a CNN with three convolutional layers (containing 32, 128, and 512 nodes respectively) is the best for projection classification. All of the best results are from the largest training size tested.

Task	**Methods**	**Testing Accuracy Rate (%)**
Map identification	MLP	95
	SVM	100
	CNN	100
Regional classification	MLP	45
	SVM	100
	CNN	100
Projection classification	MLP	40
	SVM	82.5
	CNN	85

Table 3: Results of map identification

4. Concluding Remarks

In summary, three machine learning methods are used to achieve the research objectives: map identification, region classification, and projection classification. The experiment results indicate that machine learning methods are able to identify and classify maps correctly based on region and projection. This positive result is the first step toward more applications of automatic analysis of online maps. For example, one of the future research topics is to identify map elements for automatic map evaluation and map reading.

We note that the testing size is relatively small. This is because both MLPs and CNNs need a large amount of input data to train model parameters, which makes it necessary to use a high proportion of data for training. We plan to increase the proportion of testing images to a higher percentage as more map and non-map images will be collected for future work. We will also test transfer learning using popular pre-trained CNN architectures to identify and classify maps besides the self-designed CNN architectures used in this paper. Additionally, more images will also be utilized to test the effects of training size on testing accuracy of the three methods for the three tasks.

5. List of References

Bentes, C., Velotto, D. and Lehner, S., 2015, July. Target classification in oceanographic SAR images with deep neural networks: Architecture and initial results. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3703-3706.

Chang, C.C. and Lin, C.J., 2011. LIBSVM: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, *2*(3), p.27.

Cortes, C. and Vapnik, V., 1995. Support-vector networks. *Machine learning*, *20*(3), pp.273-297.

Duda, R.O., Hart, P.E. and Stork, D.G., 2001. *Pattern Classification*. New York, NY: Wiley.

Goodfellow, L., Bengio, Y. and Courville, A., 2016. *Deep learning*. Cambridge, Massachusetts: The MIT Press.

Haykin, S., 2009. Neural Networks and Learning Machines. Upper Saddle River, NJ: Pearson education.

LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. *nature*, 521(7553), pp.436-444.

Lemley, J., Abdul-Wahid, S., Banik, D. and Andonie, R., 2016. Comparison of recent machine learning techniques for gender recognition from facial images. In *2016 Modern Artificial Intelligence and Cognitive Science Conference(MAICS)*, pp.97-102.

Qian, Y., Zhou, W., Yan, J., Li, W. and Han, L., 2015. Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery. *Remote Sensing*, 7(1), pp.153-168.

Sibanda, W. and Pretorius, P., 2011. Novel application of Multi-Layer Perceptrons (MLP) neural networks to model HIV in South Africa using Seroprevalence data from antenatal clinics. *International Journal of Computer Applications*, *35*(5), pp.26-31.

Zanaty, E.A., 2012. Support vector machines (SVMs) versus multilayer perception (MLP) in data classification. *Egyptian Informatics Journal*, *13*(3), pp.177-183.