THE OHIO STATE UNIVERSITY

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

Jialin Li and Ningchuan Xiao

Department of Geography

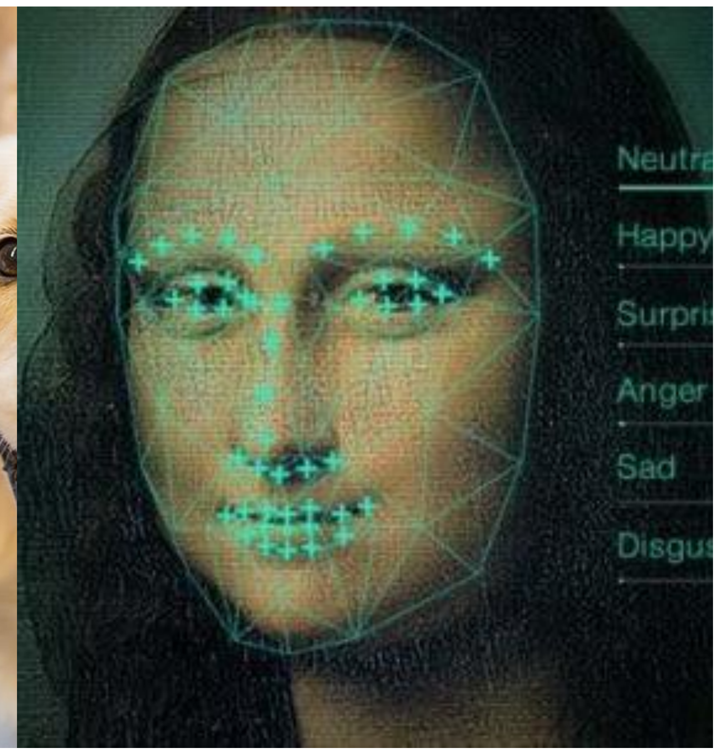
Sep 20, 2019

Introduction

- The evolution of machine learning, especially the development of convolutional neural networks, prompts new trends of machine learning applications in computer vision

Self-driving car

Image classification



Face recognition

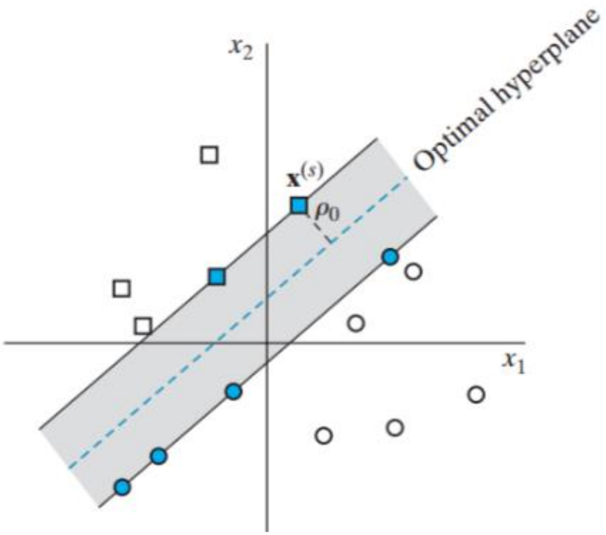
Research Questions

- Can a machine understand images such as maps?
- Research questions:
  - Is the provided image a map?
  - Which geographical region is this map about?
  - Which projection is used in this map?
- We identified and tested three methods that have potential in this research:
  - SVM (support vector machine)
  - MLP (multilayer perceptrons)
  - CNN (convolutional neural network)

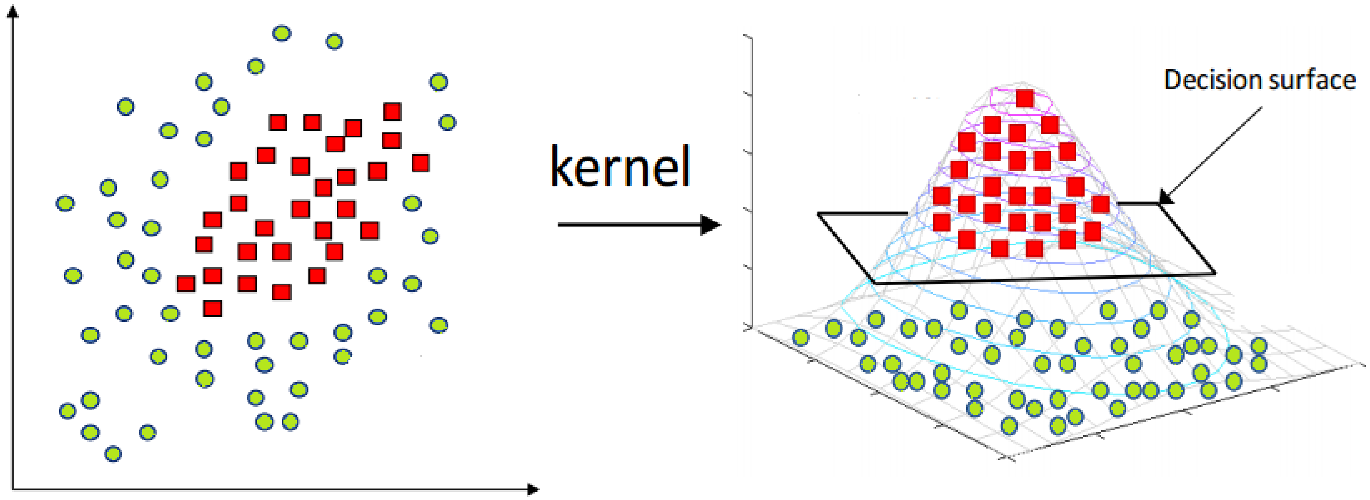
Literature Review

- Lots of relevant research in other fields including remote sensing image, medical image and plant image
  - Bentes, Velotto and Lehner (2015) use a self-designed CNN (Convolutional Neural Network) to conduct object classification in oceanographic SAR images.
  - Qian et al. in 2015 compare four machine learning classifiers for land cover classification using very high-resolution images
  - Mohanty et al. in 2016 utilize 54306 images to train AlexNet and GoogLeNet architectures in Deep CNN to identify 26 diseases from 14 distinct crop species.
- Few research about map identification and classification
  - Zhou et al. (2018) applied several state-of-the-art Deep CNN to test their performance on map-type classification. The map types include topographic map, 3D map and nighttime imagery map.

Support Vector Machine (SVM)



(Source: Haykin, 2009. Neural Networks and Learning Machines.)



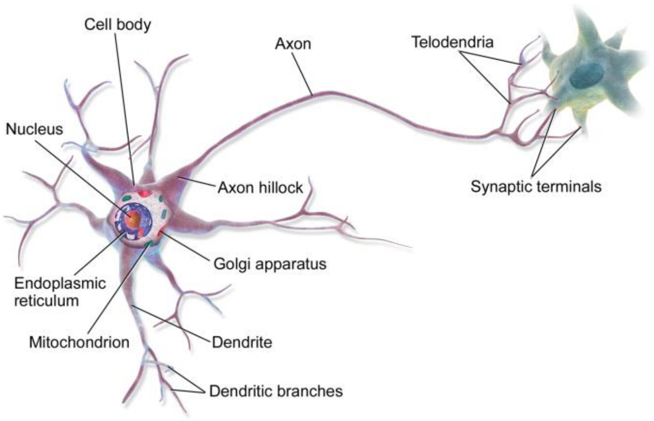
(Source: https://www.hackerearth.com/blog/developers/simple-tutorial-svm-parameter-tuning-python-r/)

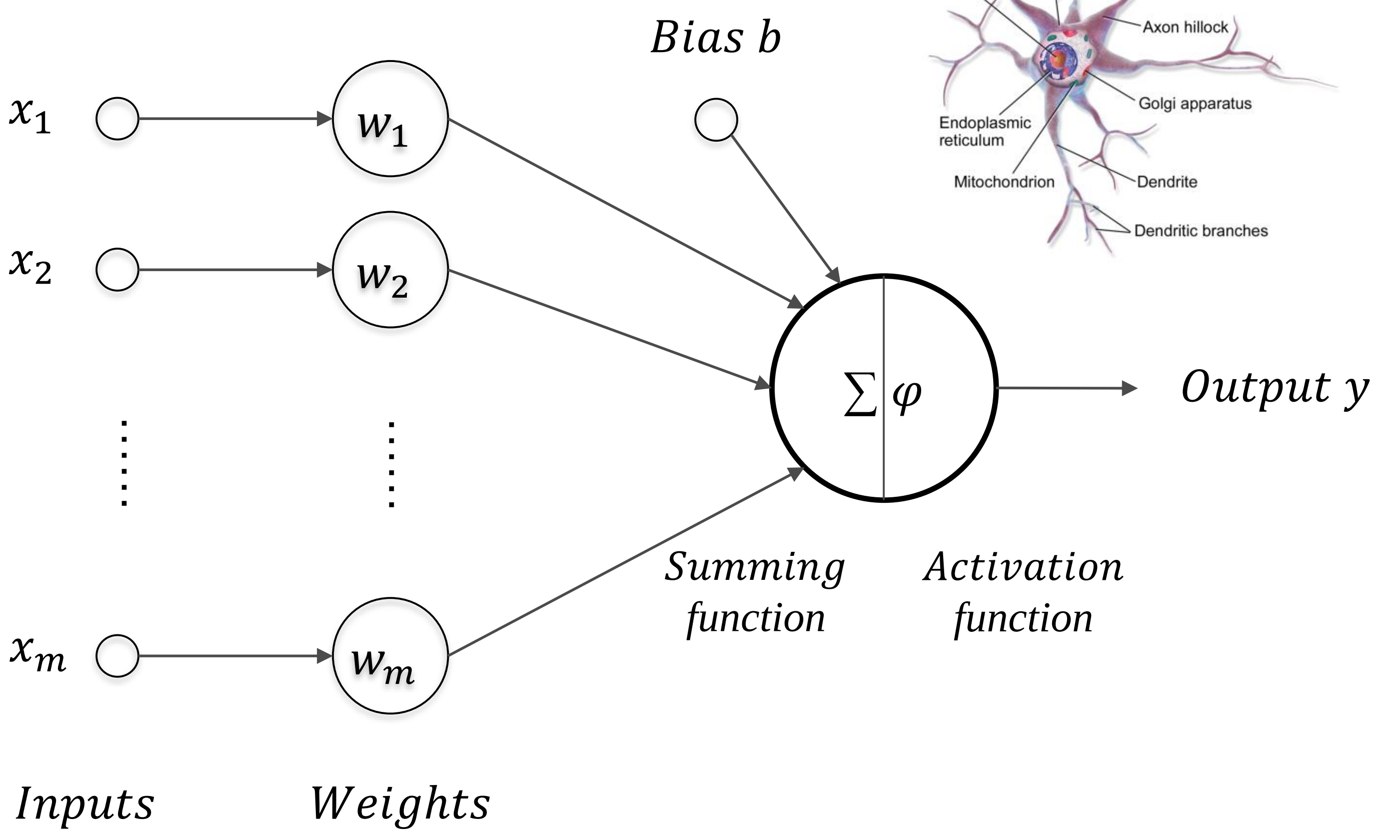
Support Vector Machine (SVM)

- Find an optimal hyperplane to separate inputs into classes
  - Optimal hyperplane: a hyperplane with a maximum margin between the classes and minimal misclassification error
  - For linearly inseparable cases, transform the original data into a higher dimensional space using kernel functions
- Four commonly-used kernel functions were tested
  - linear, polynomial, radial basis function and sigmoid kernels
- Five-fold cross validation was used to find the optimal combination of parameters

Multi-layer Perceptrons (MLPs)

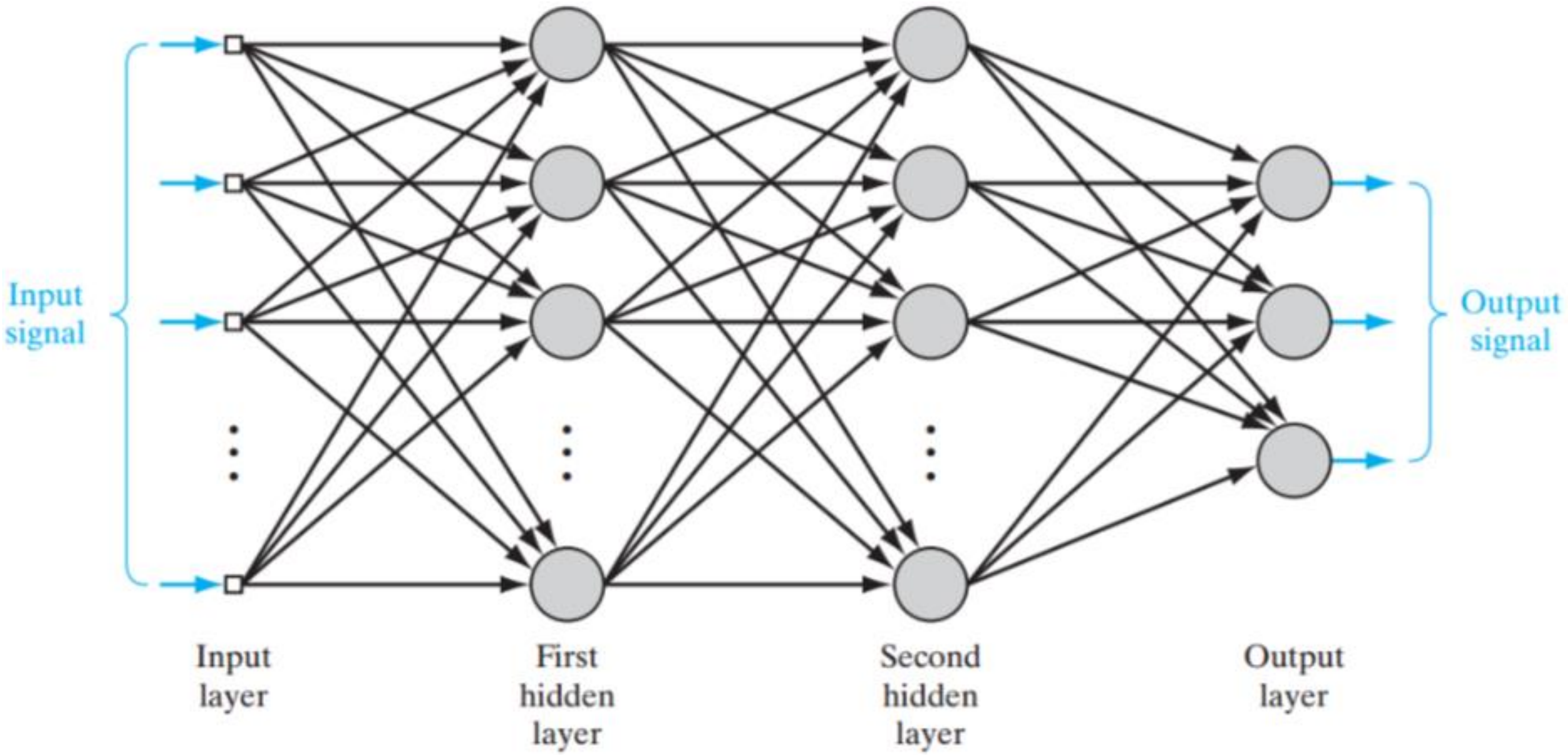
- Perceptrons are the basic elements of MLPs





Multi-layer Perceptrons (MLPs)

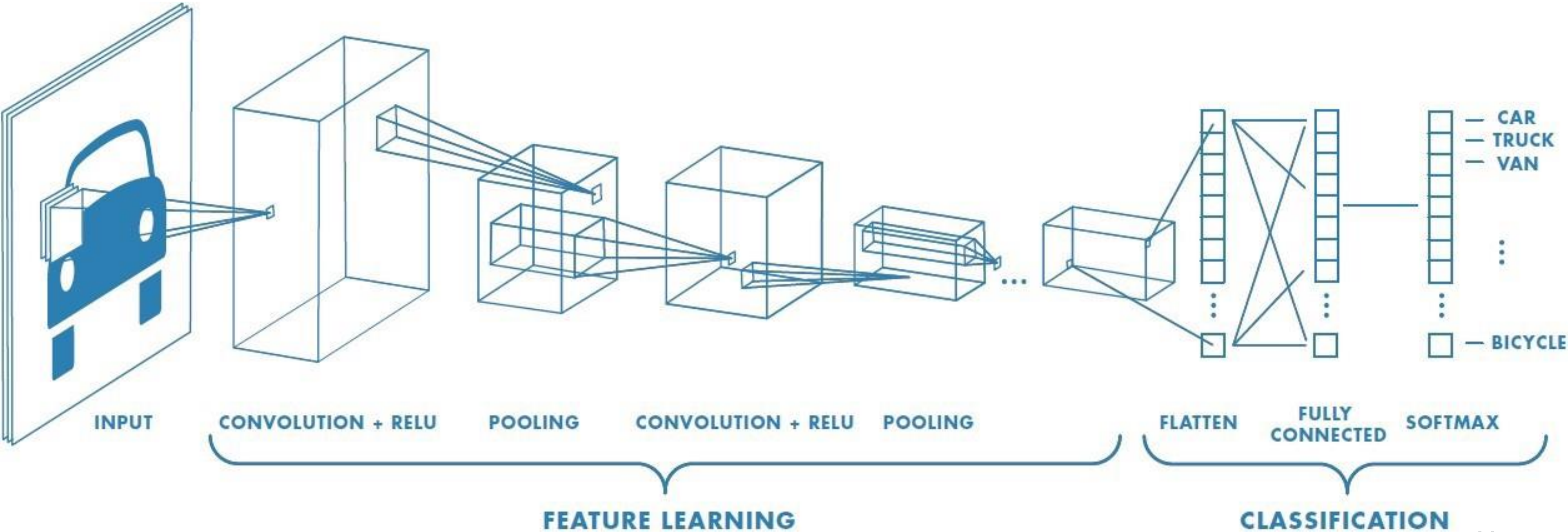
- MLPs are a type of feed-forward neural networks
- Traditional shallow networks



(Source: Haykin, 2009. Neural Networks and Learning Machines.)

Convolutional Neural Network (CNN)

- CNN can discover structures in images
- Convolutional layers are designed to discover features of different levels
- The structure of CNN can largely reduce the number of parameters and make deeper networks feasible



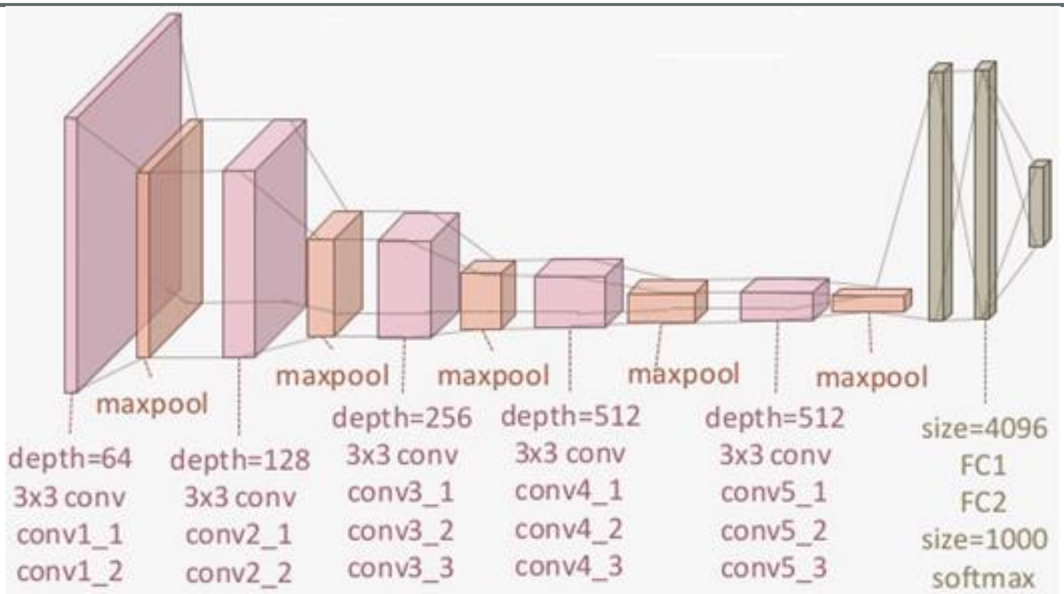
(Source: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks)

Convolutional Neural Network (CNN)

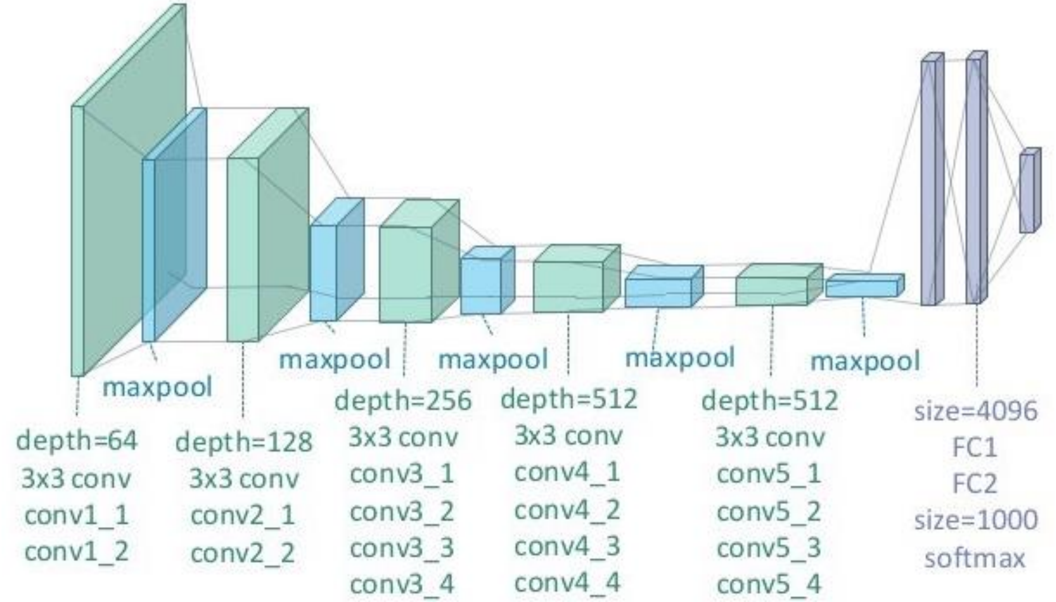
- In this study, many self-designed networks were tested
- Transfer learning using pre-trained CNN architectures was also used in the experiments
- Two commonly-used pre-trained architectures: VGG16 and VGG19 by Visual Geometry Group from Oxford
  - The VGG network architecture was introduced by Simonyan and Zisserman in 2014
  - VGG networks are one of the most commonly-used pre-trained architectures
  - The “16” and “19” stand for the number of weight layers

Convolutional Neural Network (CNN)

VGG16 Architecture



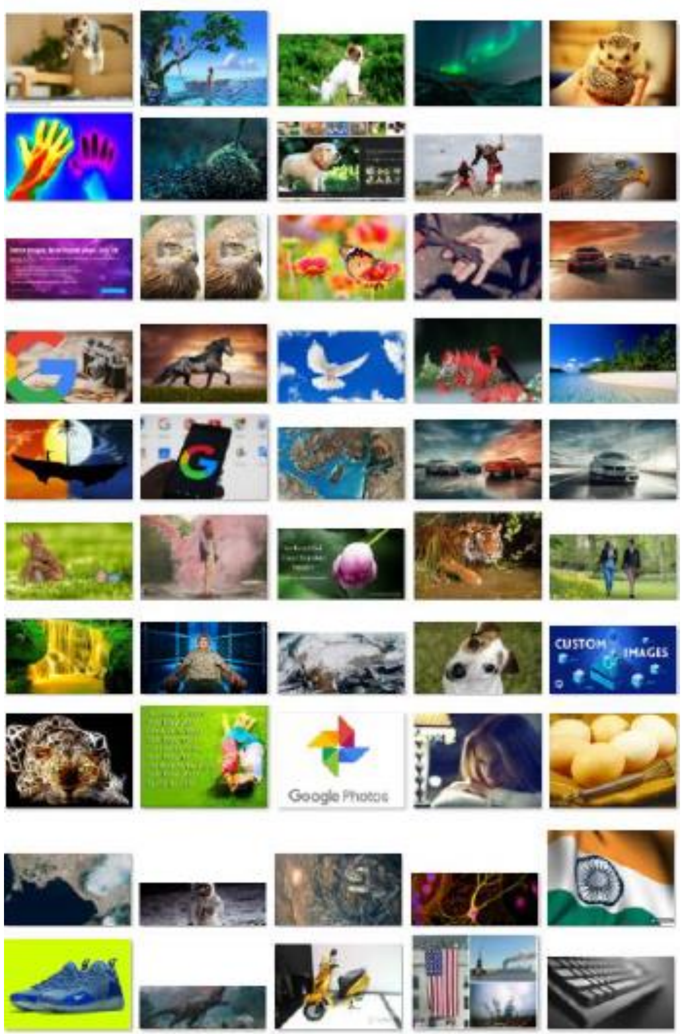
VGG19 Architecture



(Source: https://towardsdatascience.com/transfer-learning-in-tensorflow)

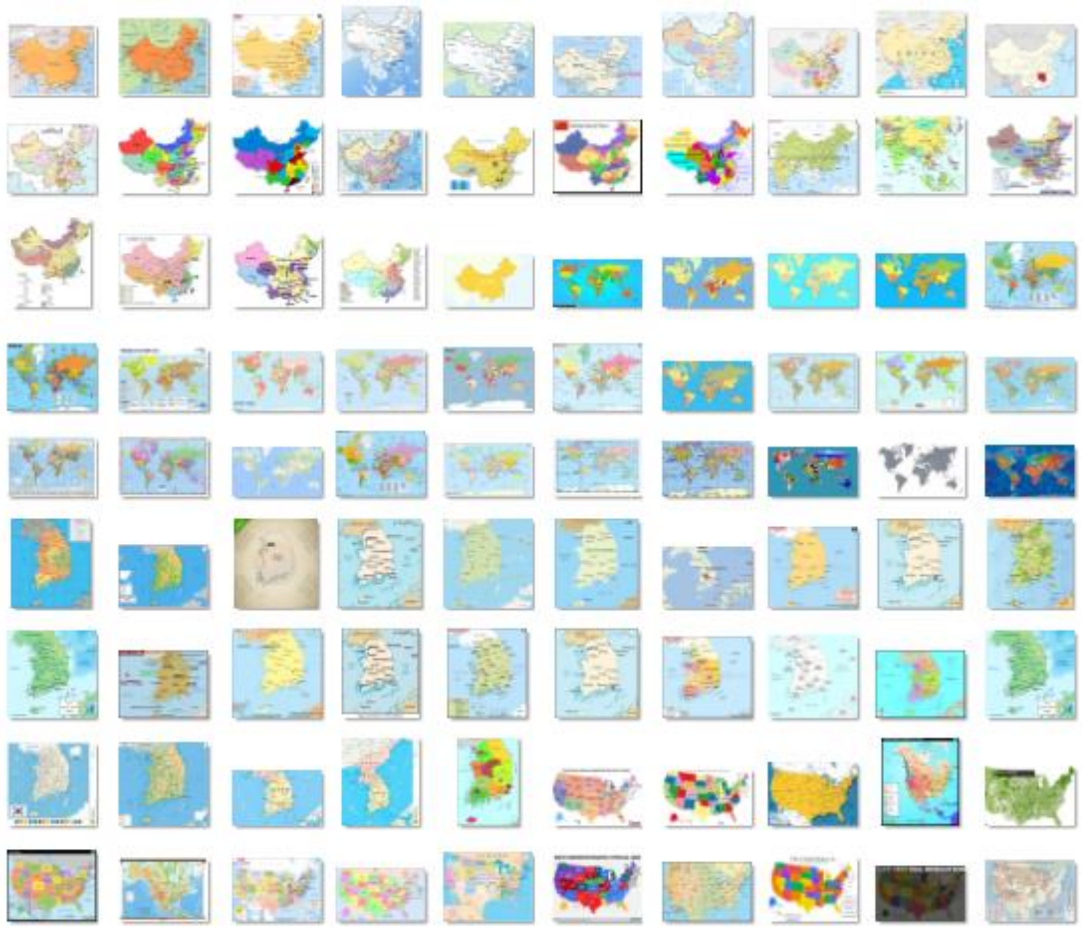
Data: Map Identification

- 500 non-map images and 500 world map images



Data: Region Classification

- 250 maps for each of four regions (China, the U.S., South Korea and the World)



Data: Projection Classification

- 100 map images for each of four world map projections (Equirectangular, Mercator, Miller and Robinson projection)

Tools

- TensorFlow
  - Open-source machine learning library by Google
- Keras
  - Runs on top of TensorFlow
  - Designed to enable fast implementation
  - Used for MLPs and CNNs
- LIBSVM
  - A popular open-source SVM library to implement SVMs

Computational Experiments

- For map identification and region classification, 70% images are training data, and the rest 30% are testing data
- For projection classification, 90% images are training data and 10% are testing data
- The inputs for all three methods are pixel values in three bands of each image (3*120*100 input nodes)
- For SVMs, MLPs and CNNs trained from scratch, only CPU (Intel Core i7-8700, 3.2GHz) was used to run the program
- For transfer learning using pre-trained CNN architecture, the program ran on Google Colab with GPU (Nvidia Tesla K80, 2496 CUDA cores, 12 GB memory)
  - Google Colab is a Jupyter notebook environment that requires no setup to use and runs entirely on the cloud

Hyper-parameters

MLPs

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

CNNs

Parameters	Values or Choices
Activation function	Rectified linear unit (ReLU)
Kernel size	5*5
Pool Size	2*2
Strides	2*2
Optimizer	Stochastic gradient descent
Learning rate	0.01

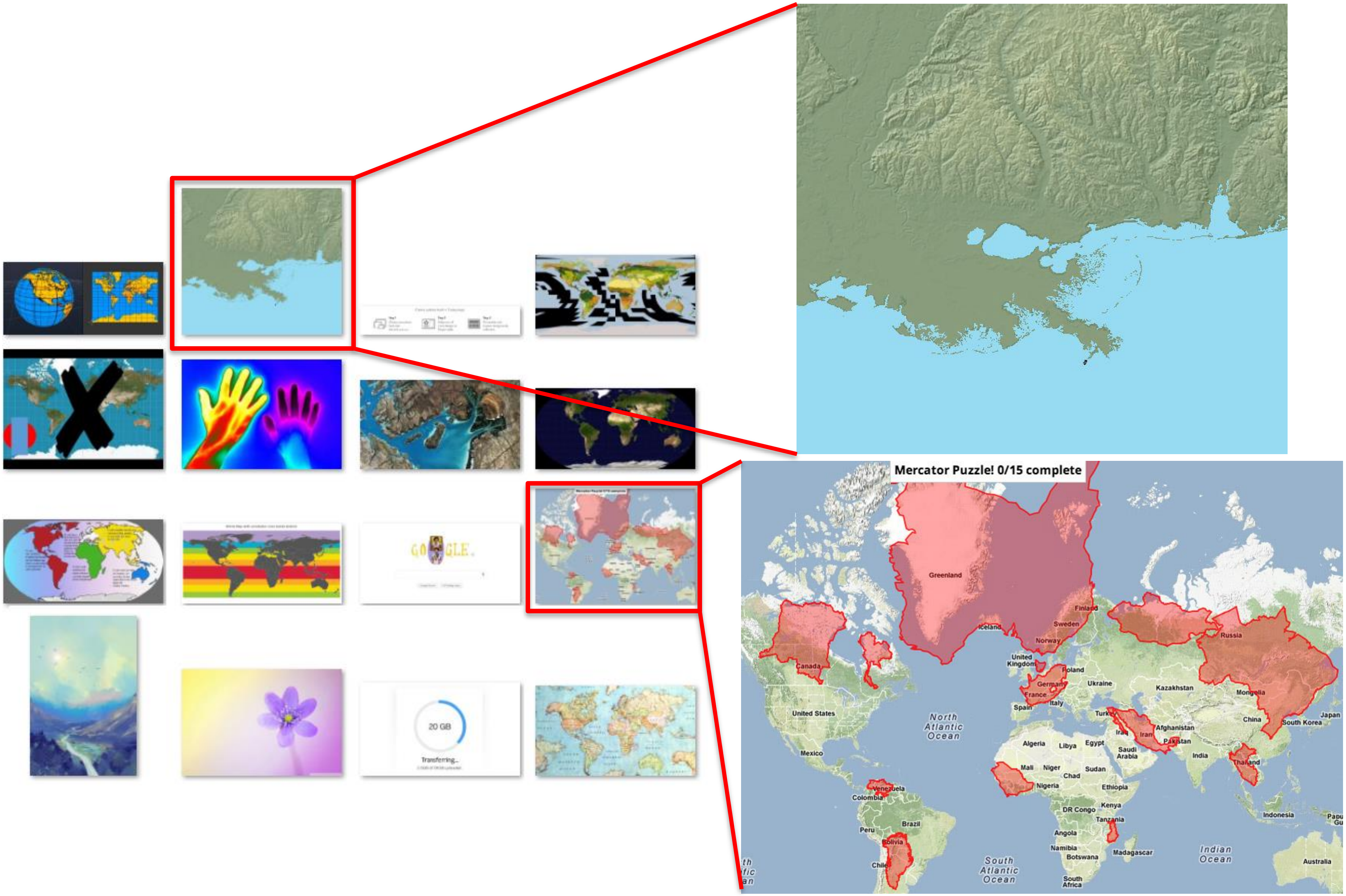
Results: Map Identification

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$

Methods	Training Accuracy	Testing Accuracy
MLP	0.931	0.903
SVM	0.998	0.857
CNN	0.997	0.950

- The best result is from CNN with three convolutional layers, and the kernel numbers of each layers are 16, 64, 256
- All of the three methods can get high testing accuracy, and CNN is the best in general and also the most stable method

Results: Incorrect Predictions



Results: Region Classification

Methods	Training Accuracy	Testing Accuracy
MLP	0.764	0.784
SVM	0.998	0.790
CNN	0.999	0.837

- The testing accuracy rates of the three methods are not as good as those for map identification
- The best result is from CNN with two convolutional layers, and the kernel numbers of each layers are 32 and 64
- CNN's training accuracy is much higher than testing accuracy

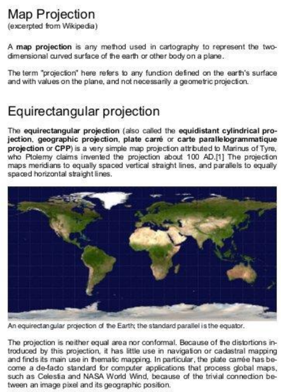
Results: Incorrect Predictions

Results: Projection Classification

Methods	Training Accuracy	Testing Accuracy
MLP	0.764	0.775
SVM	0.998	0.85
CNN	0.999	0.9

- The testing accuracy rates of the three methods are not as good as those for map identification
- The best result is also from CNN with two convolutional layers, and the kernel numbers of each layers are 64 and 128

Results: Incorrect Predictions



Map Projection
(excerpted from Wikipedia)

A **map projection** is any method used in cartography to represent the two-dimensional curved surface of the earth or other body on a plane.

The term "projection" here refers to any function defined on the earth's surface and with values on the plane, and not necessarily a geometric projection.

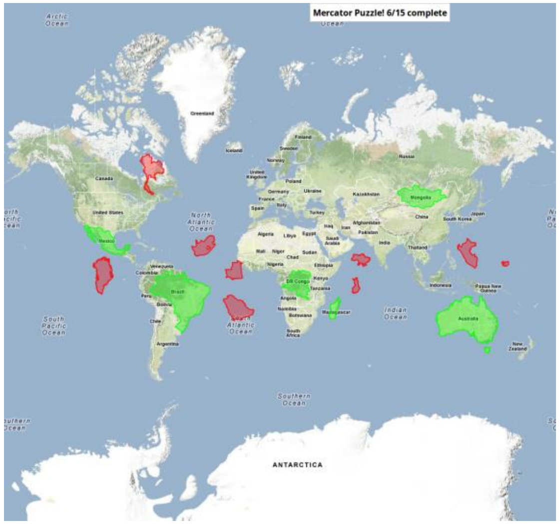
Equirectangular projection

The **equirectangular projection** (also called the **equidistant cylindrical projection**, **geographic projection**, **plate carré** or **carte parallelogrammatique projection** or **CPP**) is a very simple map projection attributed to Marinus of Tyre, who Ptolemy claims invented the projection about 100 AD.[1] The projection maps meridians to equally spaced vertical straight lines, and parallels to equally spaced horizontal straight lines.

An equirectangular projection of the Earth; the standard parallel is the equator.

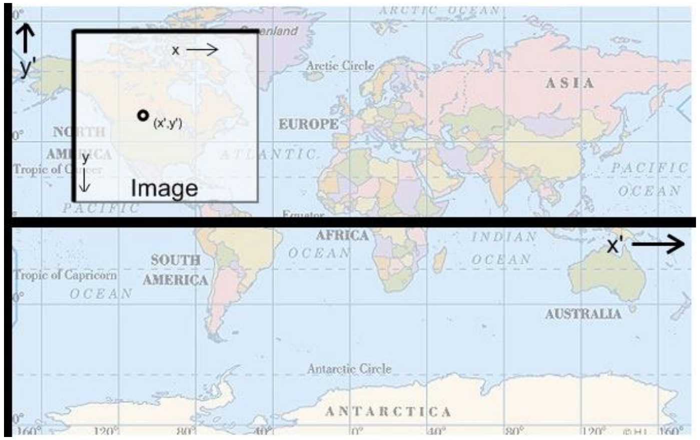
The projection is neither equal area nor conformal. Because of the distortions introduced by this projection, it has little use in navigation or cadastral mapping and finds its main use in thematic mapping. In particular, the plate carrée has become a de-facto standard for computer applications that process global maps, such as Celestia and NASA World Wind, because of the trivial connection between an image pixel and its geographic position.

Equirectangular - Robinson



Mercator - Miller

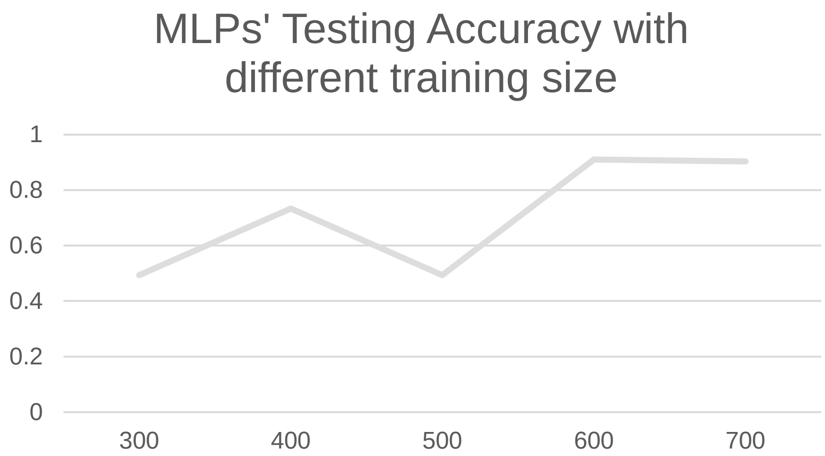
Miller - Mercator

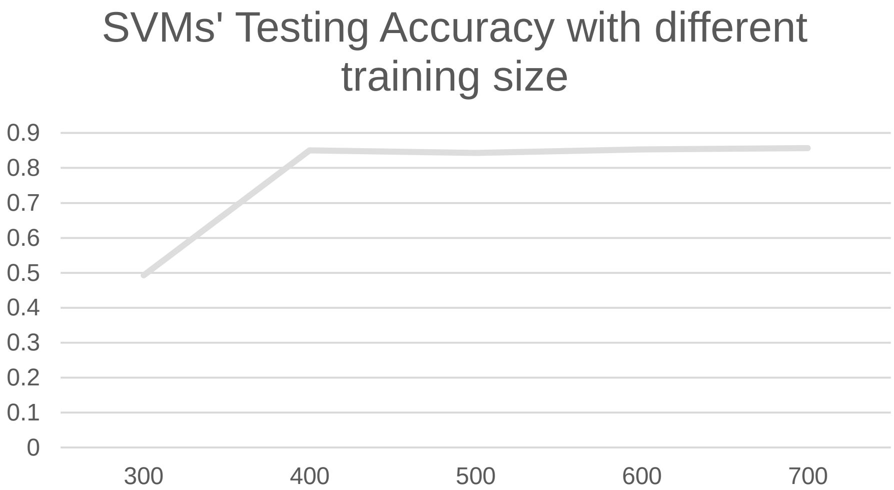


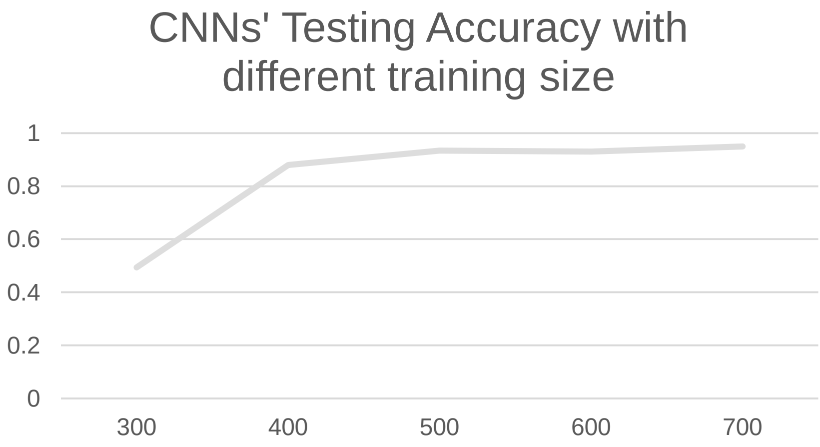
Miller - Mercator

Results: Training Size

- Models with more training data will obtain higher accuracy rate
- Choose map identification as an example







Results: Transfer Learning

Task	Methods	Training Accuracy	Testing Accuracy
Map identification	CNN from scratch	0.997	0.950
	VGG16	0.989	0.987
	VGG19	0.999	0.987
Region classification	CNN from scratch	0.999	0.837
	VGG16	0.997	0.933
	VGG19	0.997	0.947
Projection classification	CNN from scratch	0.999	0.9
	VGG16	0.972	0.675
	VGG19	0.972	0.725

Results: Training and Testing Time

Task	Methods	Training Time (s)	Testing Time (s)
Map identification	CNN from scratch	1749.42	0.69
	VGG16	901.54	4.36
	VGG19	1101.46	5.87
Regional classification	CNN from scratch	2618.63	0.60
	VGG16	528.97	3.51
	VGG19	1136.71	6.23
Projection classification	CNN from scratch	362.56	0.11
	VGG16	203.50	0.67
	VGG19	255.67	0.85

Summary

- Significance of our research work:
  - Basic research field in AI and machine learning
  - Fundamental work of automatic map understanding
  - A good assistant for blind or vision-impaired users to extract information from map images
  - Useful for image retrieval and update in an image database
- Future work:
  - Try more state-of-the-art pretrained CNN architectures
  - Use high-performance computation techniques to reduce computing time