Learning Educational Concepts from Images of Pinterest with Artificial Intelligence

Bin Hwang

Xinyu Lei

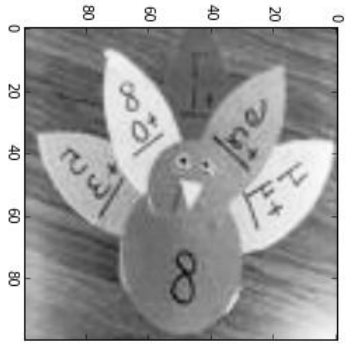
Zhiwei Wang

Overview

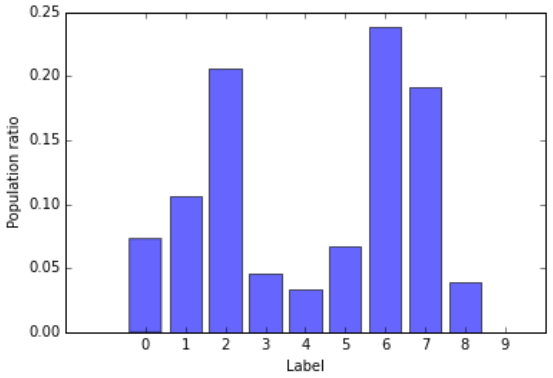
- Data: Training data set is available from the researchers in College of Education of MSU.
- Labels: 9 different labels including ‘Art’, ‘Concrete manipulatives’, ‘Visual representation’..
- Preprocessing: Clean duplicate images, Renormalized, Rescaled, Data Balance, PCA
- Methods: ANN, DNN written by Python scripts only using numpy, CNN with TensorFlow
- Loss function : Square loss for ANN/DNN , Cross entropy loss for CNN

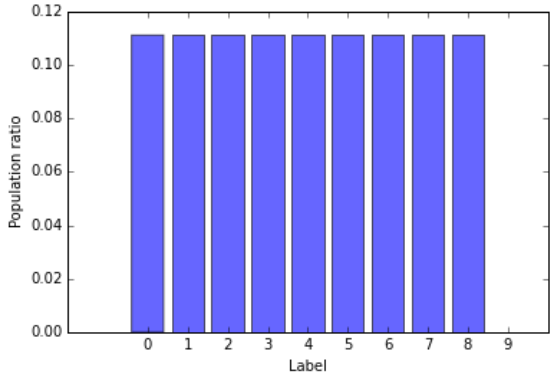
Data Preprocessing

- Resize all images to 32x32, 100x100 and color/gray scale



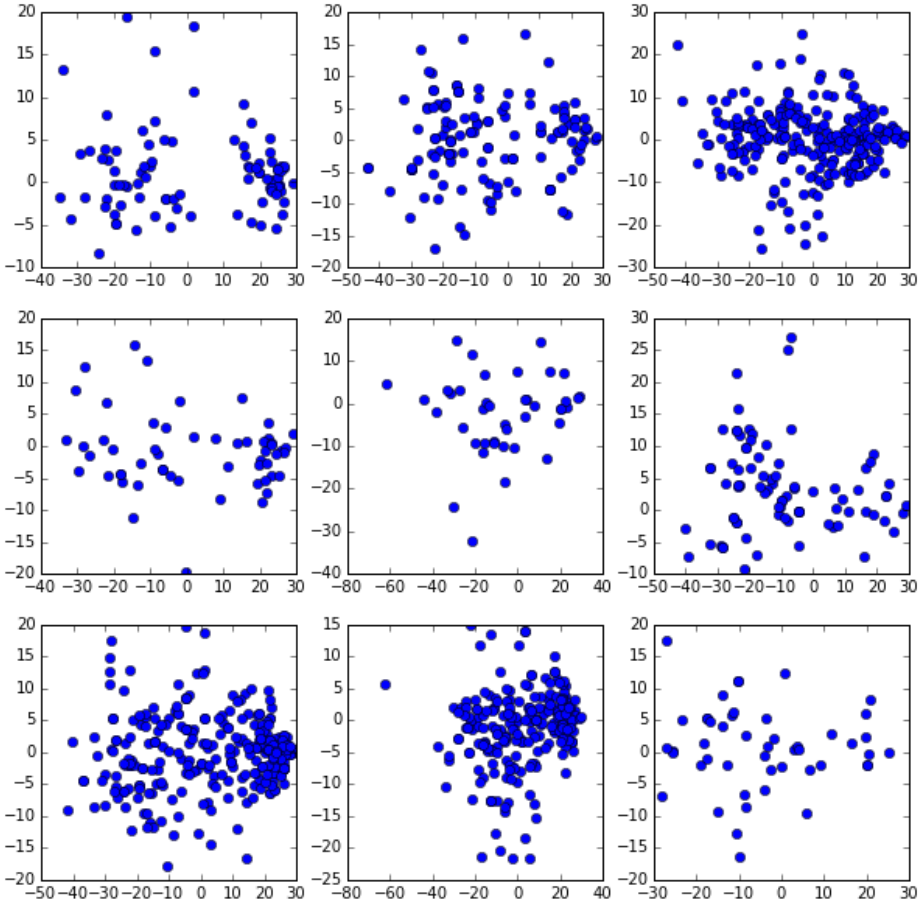
- Rebalance the data with duplicating small data set from 1200 figures to 2250 figures with 80% training data (1800 figures) and 20% testing (450 figures)





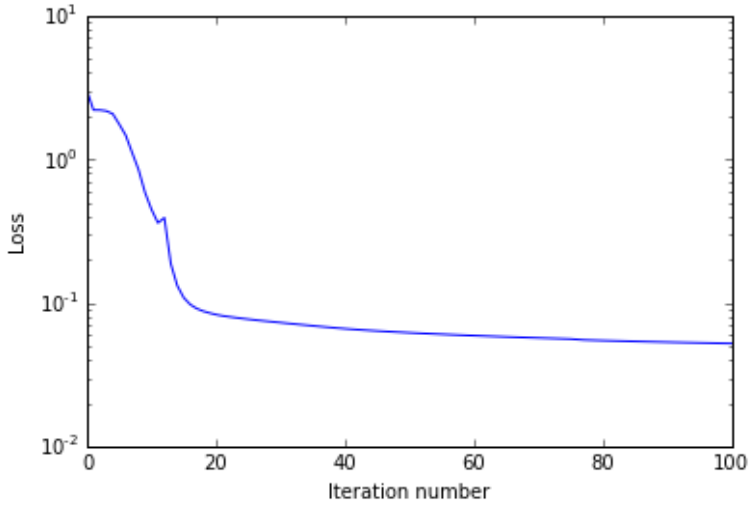
Principle Component Analysis (PCA)

- Understand data with a rough picture
- Use 2-component PCA to get some information

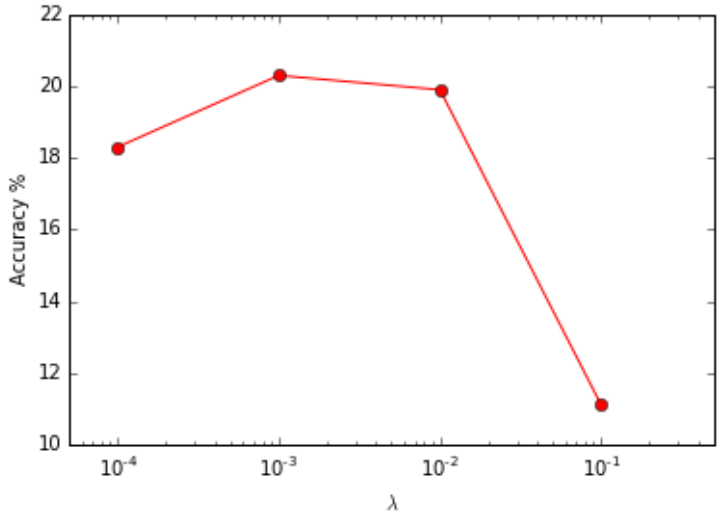


ANN/DNN with different regression

- Square Loss : (y-f(X))^2
- Loss monotonically decrease for most of iterations



- Regression Loss: λ $(w-w_0)^2$
- Regression coefficient with DNN layer size [20,10]

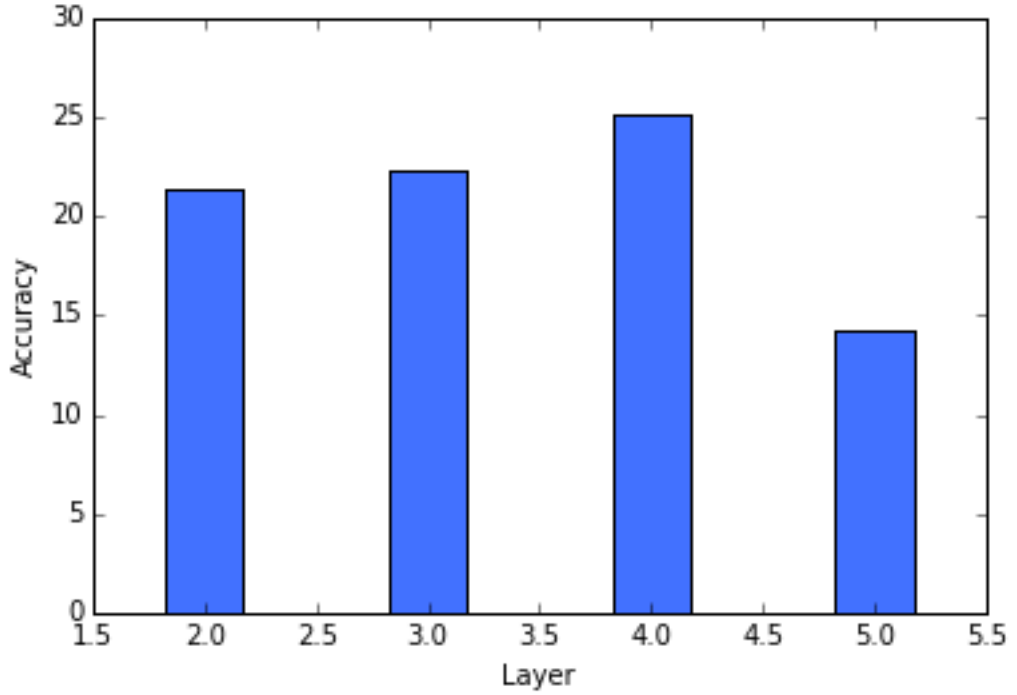


- Best Accuracy :21.23% with λ := 0.001
- Baseline Accuracy : 1/9 = 11.11%

ANN/DNN with different layers

- ANN : 2 layers with layer size [10]
- DNN: 3 layers with layer size [20,10], 4 layers with layer size [30,20,10],

  5 layers with layer size [40,30,20,10]
- All cases use λ := 0.001
- Best Accuracy: 25.56%



- Confusion matrix:

[5 13 3 3 0 0 12 15 0]
[0 18 12 1 4 2 10 4 0]
[3 6 12 2 3 1 12 12 0]
[0 5 19 0 5 0 8 14 0]
[5 5 9 0 10 4 0 14 4]
[6 7 12 0 0 9 14 0 3]
[2 7 11 2 0 3 17 9 0]
[0 6 9 3 1 3 10 18 1]
[0 5 16 0 0 0 0 25 5]

CNN-Based Method

- Platform: Tensorflow and Python.

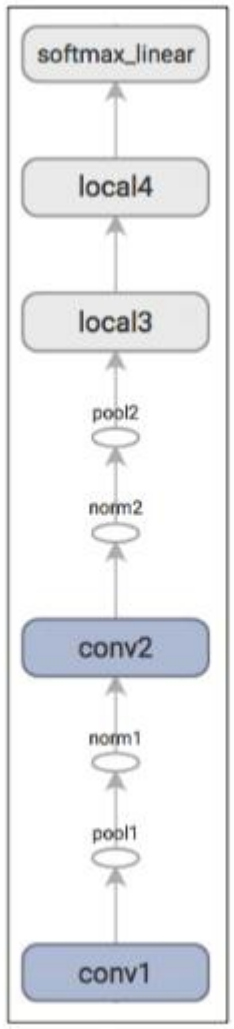
- Introduction: TensorFlow™ is an open source software library from **Google** for numerical computation using data flow graphs. TensorFlow has better computational **graph visualization** and clear **data flow graph.**

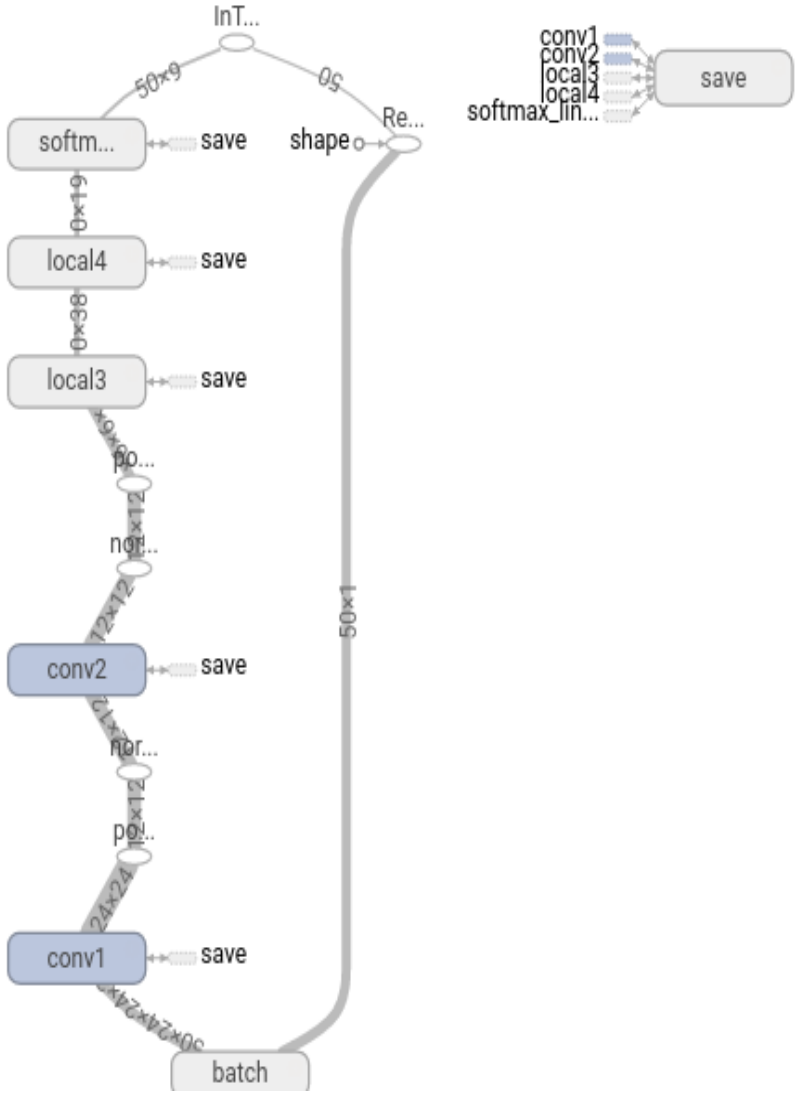
- Use the similar CNN Structure used in Cifar10 project

	1	2	3	4	5	6	7	8	9	10
Cifar10	airplane	Automobile	bird	cat	deer	dog	frog	horse	ship	truck
Pinterest	Art	concrete	content	contextual	Kinestetic	Peda	physi	Standard_alg	Visual_repre	

CNN Model Structure

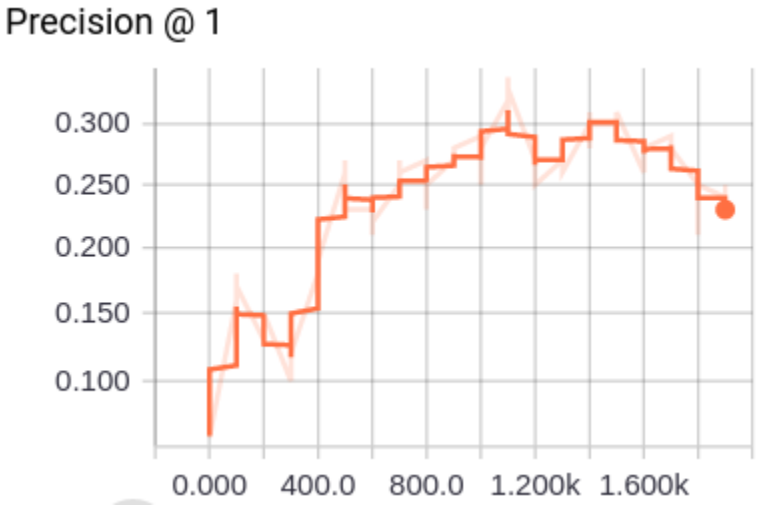
Layer Name	Description
cov1	Convolution and rectified linear activation
pool1	Max pooling
norm1	Local response normalization
conv2	Convolution and rectified linear activation
norm2	Local response normalization
Pool2	Max pooling
local3	Full connected with rectified linear activation
local4	Full connected with rectified linear activation
Softmax_linear	Linear transformation to product logits



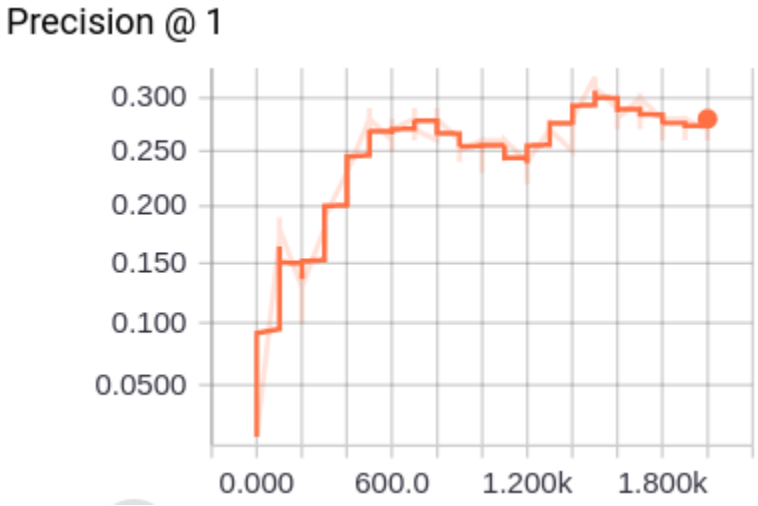


Increase Input Image dimensionality

50x50 images



75x75 images



- We can increase the precision to be about 30% by increasing the input image Dimensionality.
- Further improvement is possible, but it will take a huge amount time for training on a normal PC.

Discussion

- Question : How human understand concepts from figures ?
- Our Answer: Start from concrete materials (dogs, cats, pens…) and symbols (words and numbers) in the pictures. We can do this by putting in new features or using multi-labels.

- Re-label 1200 figures into two classes: with symbols (words and numbers) and without symbols
- 1000 training and 200 testing data
- Accuracy : 78.56% (DNN) and 80.12% (CNN) with baseline accuracy : 50%