Hand Written Digit Recognition

By Ananda Kishore Sirivella

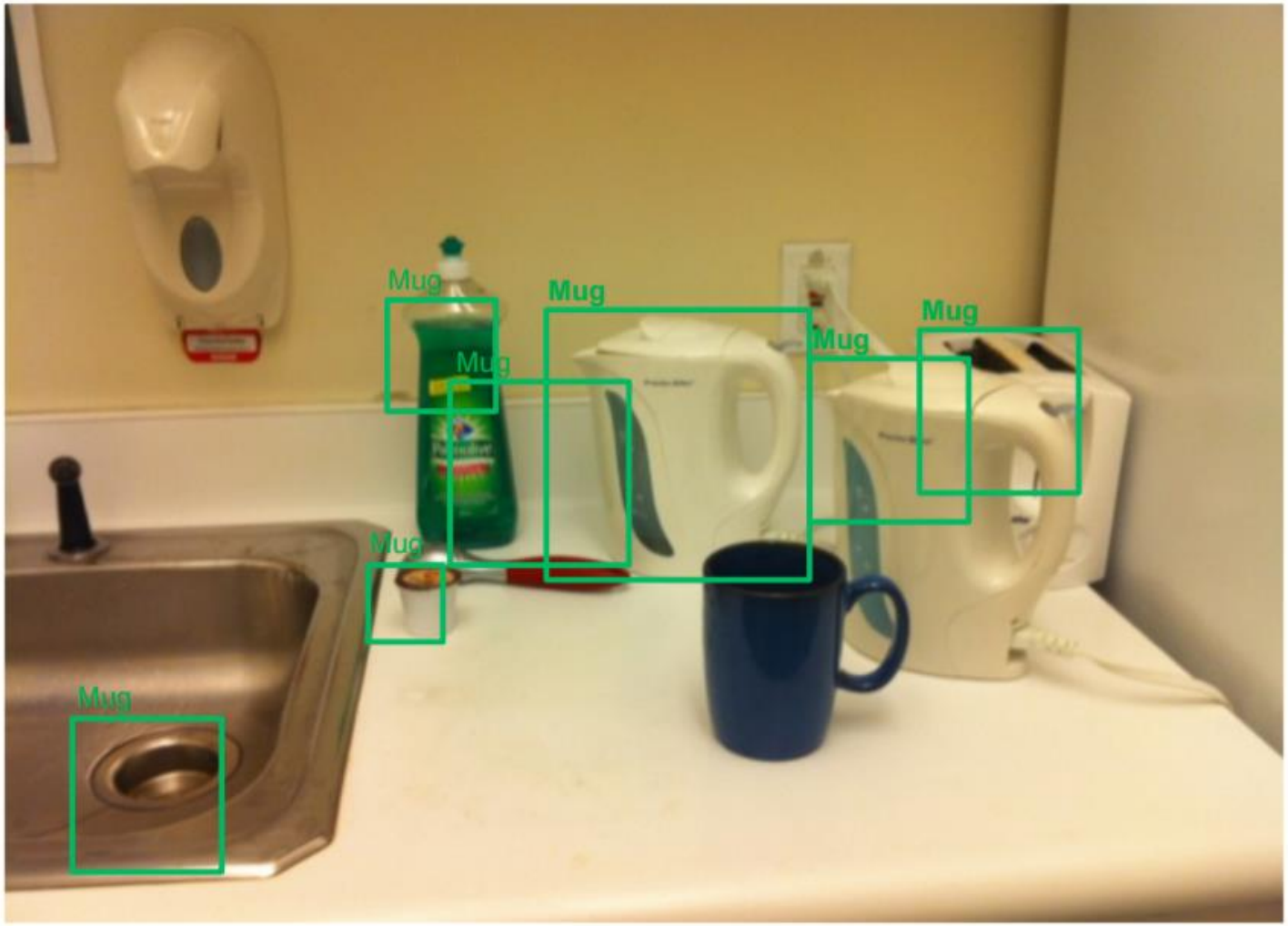
UFID: 9951-5080, EEE 6512 Fall 2017

University of Florida, Gainesville, FL

Introduction

- Developments in the field of Artificial Intelligence & Emphasis over image recognition.
- For AI to work similar to human brain, we need:-
  - Image recognition
  - Pattern recognition
  - Speech recognition
  - And many more
- The field of manipulating, recognizing and learning a image is computer vision.

Problem with computer vision



What a machine sees!

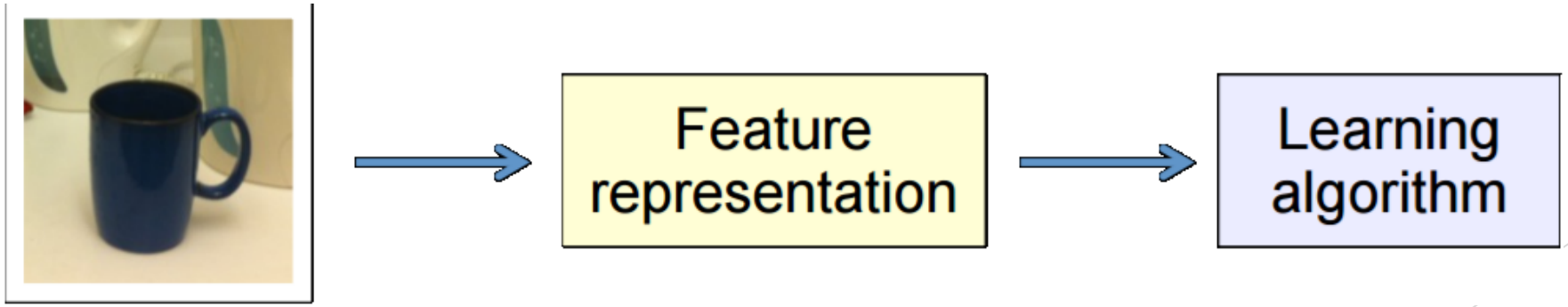
The camera sees :

194	210	201	212	199	213	215	195	178	158	182	209
180	189	190	221	209	205	191	167	147	115	129	163
114	126	140	188	176	165	152	140	170	106	78	88
87	103	115	154	143	142	149	153	173	101	57	57
102	112	106	131	122	138	152	147	128	84	58	66
94	95	79	104	105	124	129	113	107	87	69	67
68	71	69	98	89	92	98	95	89	88	76	67
41	56	68	99	63	45	60	82	58	76	75	65
20	43	69	75	56	41	51	73	55	70	63	44
50	50	57	69	75	75	73	74	53	68	59	37
72	59	53	66	84	92	84	74	57	72	63	42
67	61	58	65	75	78	76	73	59	75	69	50

Proposed Solution

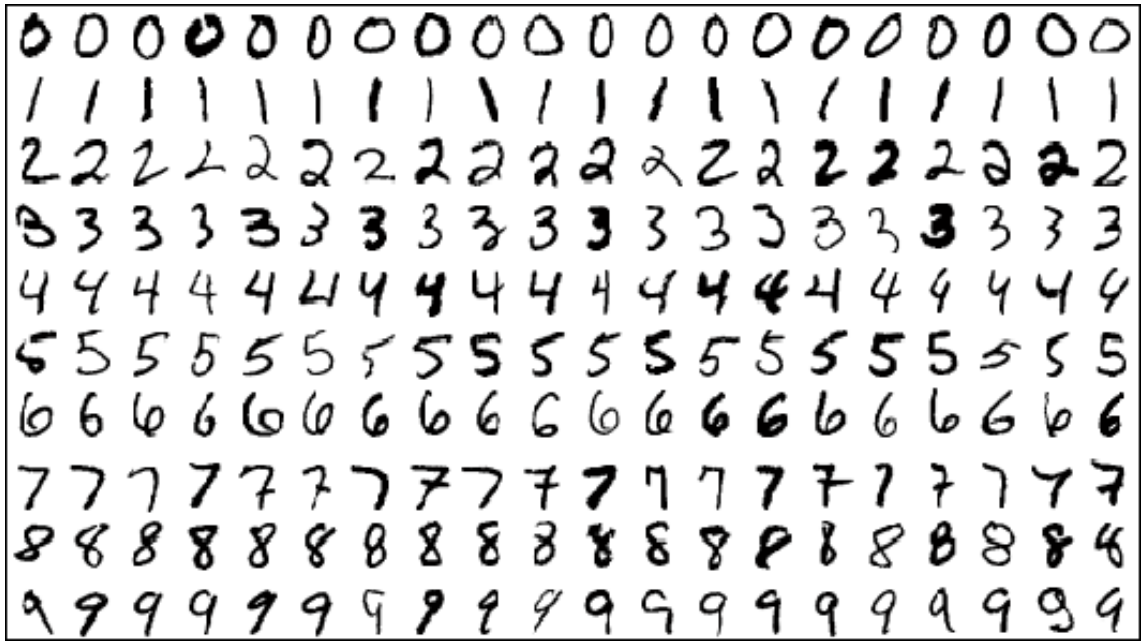
Machine learning being latest science to such algorithms, where algorithms are developed to learn from the data and re-construct is model to provide better accuracy. Machine learning algorithms construct a black box such that on a input it would recognize and formulate the output.

- Supervised learning: Construction of a learning algorithm over a labeled data & tuning over test data to increase accuracy
- Un-Supervised learning: Construction of a learning algorithm over real time data. More of self-learning.



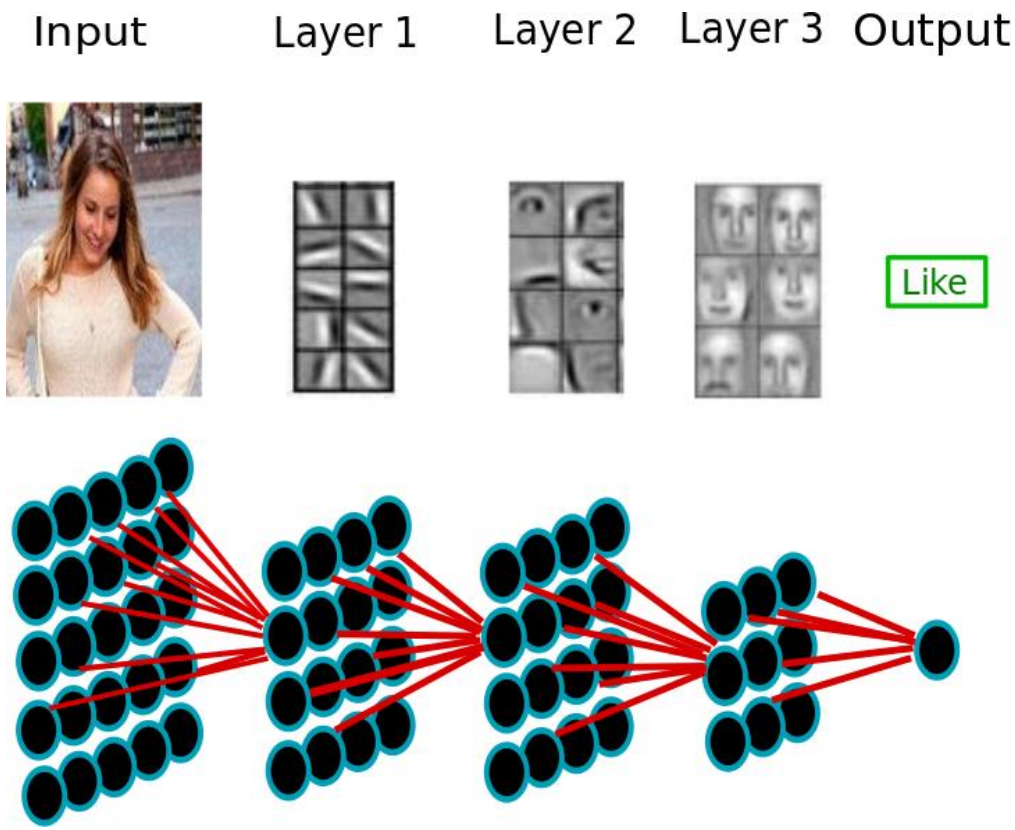
MNIST dataset

- A subset of NIST data, that is clean and normalized data.
- 60,000 samples to train and 10,000 samples to test on.
- Each image of 28 by 28 pixels .i.e. 784 pixels.



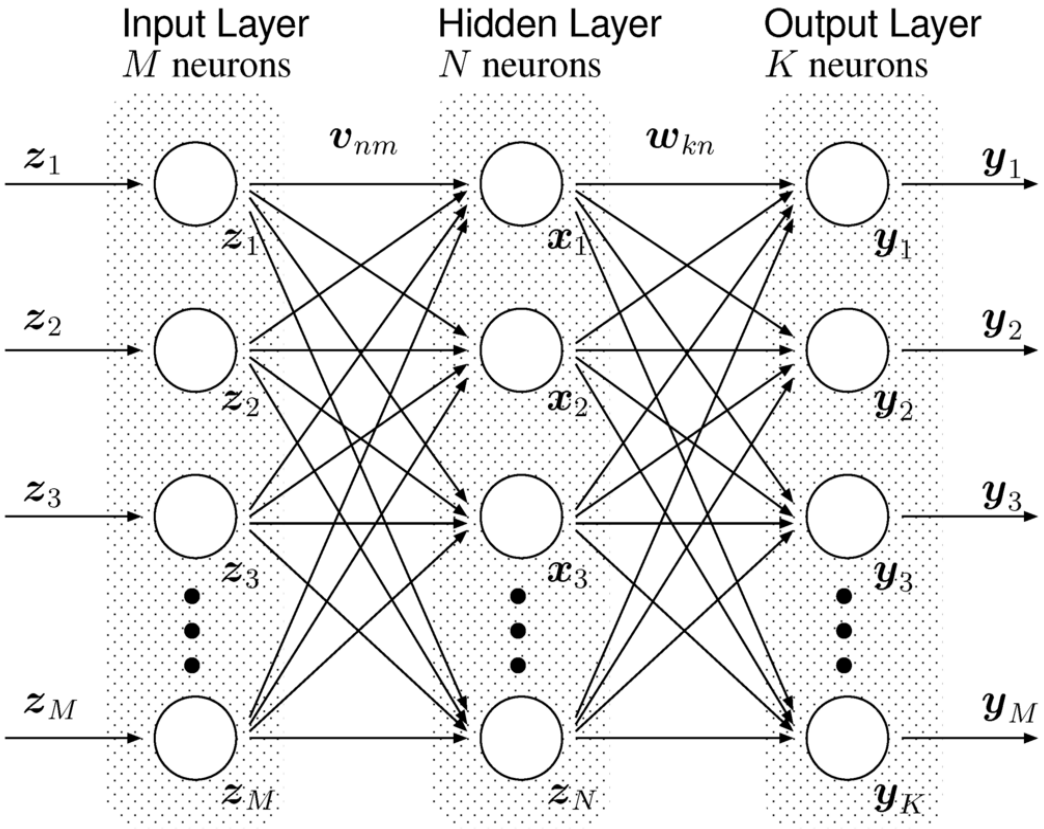
Deep Learning

- A neural network approach to the problem.



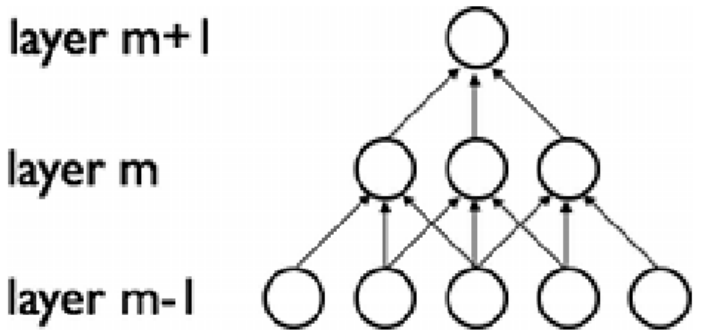
Multi-layer Perceptron

- Generally contains multiple Hidden layers.
- Utilized modified version of linear perception of data.
- Has constant Learning Rate.
- Very fast when we have good margin for error rate.



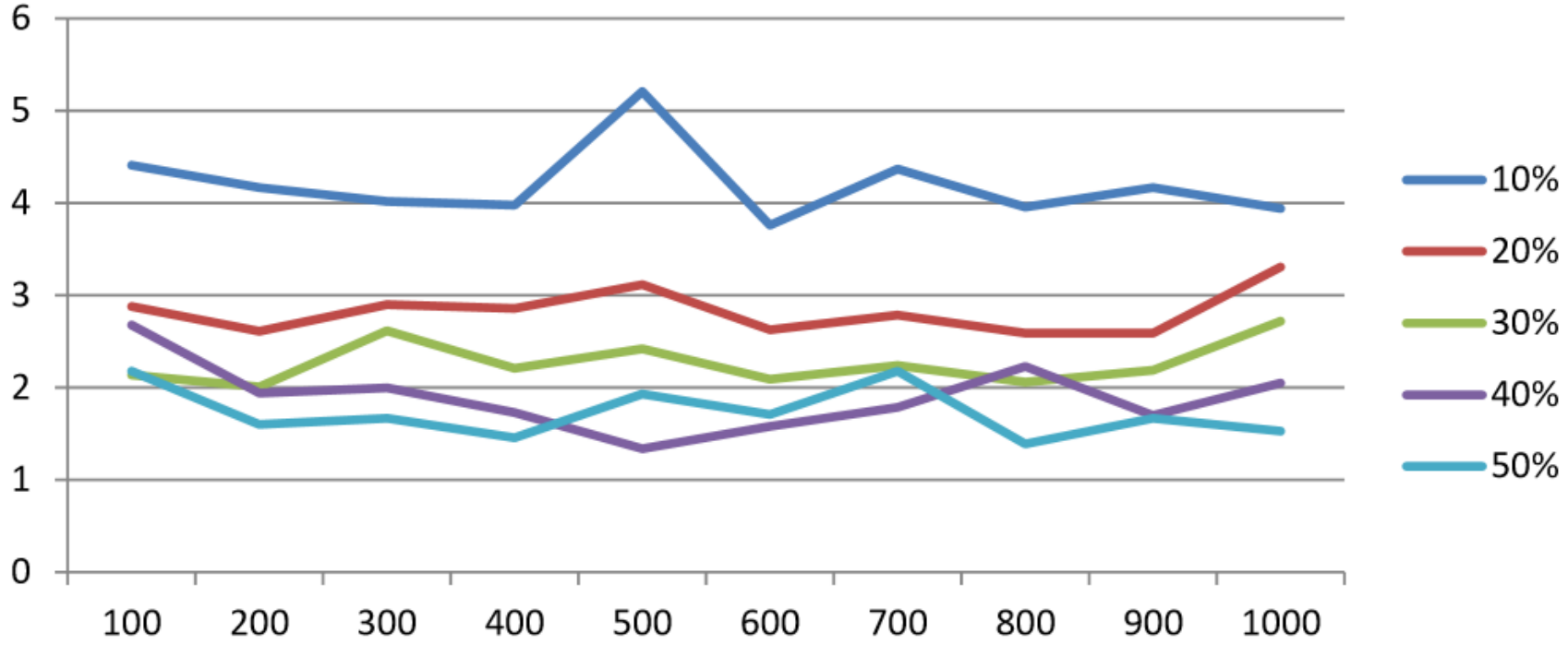
Convolutional Neural Networks

- A version of feed forward network. Better in handling in non-linear images.
- Each hidden layer has a convolutional layer and a pooling layer to improve the efficiency
- The efficiency increases over number of layers.
- But execution time multiplies over each layer added.



CNN over UCI Optical data set

- Output with 10%, 20%, 30%, 40% & 50% training set vs different values of error rate



Results on MNIST

- Result with Training set of 60,000 samples and test set of 10,000 samplese

Model	Execution time	Accuracy	Error rate
Baseline Multi-layer Perceptron	81 sec	98.02 %	1.92 %
Simple CNN	1,094 sec	99.00 %	1.00 %
Complex CNN	2,098 sec	99.21 %	0.79 %

Implementation

Implemented using some of the python libraries

- Theano:
  - Excellent framework that utilized GPU to visualize multi-dimensional arrays.
- NumPy:
  - Framework helpful for math part like fourier transform, linear algebra, trigonometry function and random functions
- Keras:
  - Runs over Theano, very much helpful, fast and efficient in implementation of Neural networks.

Conclusion

- Multi-Layer Perceptron
  - Less Accurate (~2 %) but very fast (81 seconds)
- Simple CNN
  - Better accuracy than MLP but a bit slow (~1,000 seconds)
- Complex CNN
  - Highly Accurate with increase in layer but very slow (2,096 seconds)

Reference

- keras.io
- python.org
- Google.com
- en.wikipedia .org
- yann.lecun.com/exdb/mnist/
- machinelearningmastery.com
- deeplearning.net/software/theano/
- neuralnetworksanddeeplearning.com