



FALCON AGENCY  Oduriana

The FALCON AGENCY logo consists of the word "FALCON" in a dark blue, outlined font, followed by "AGENCY" in a smaller, solid dark blue font. To the right is the Oduriana logo, which features a stylized orange and yellow circular icon followed by the word "Oduriana" in a script font.

Introduction to Deep Learning

Poo Kuan Hoong

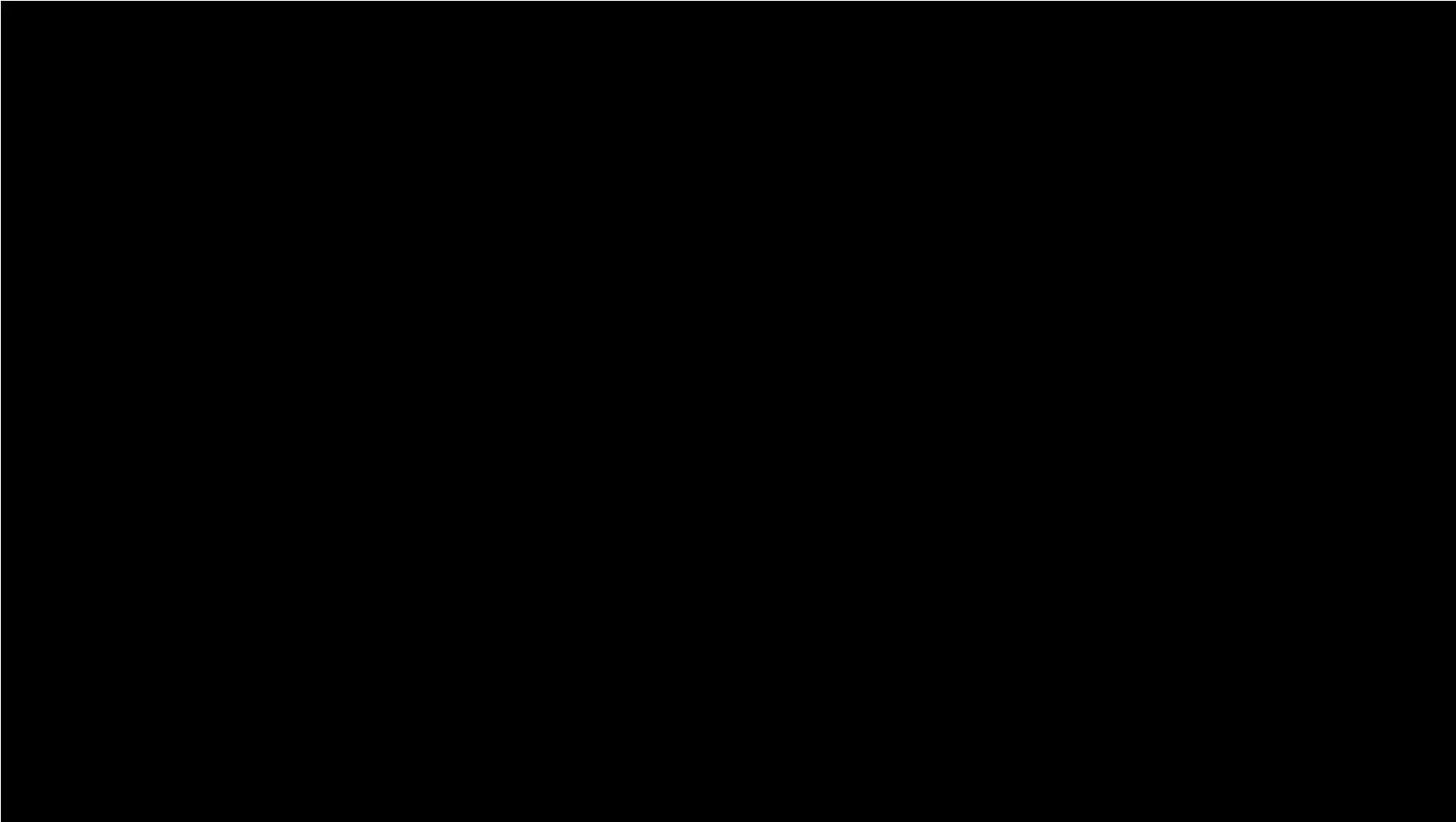
19th July 2016

Data Science Institute

- The Data Science Institute is a research center based in the Faculty of Computing & Informatics, Multimedia University.
- The members comprise of expertise across faculties such as Faculty of Computing and Informatics, Faculty of Engineering, Faculty of Management & Faculty of Information Science and Technology.
- Conduct research in leading data science areas including stream mining, video analytics, machine learning, deep learning, next generation data visualization and advanced data modelling.



Google DeepMind playing Atari Breakout



<https://www.youtube.com/watch?v=V1eYniJ0Rnk>

Google DeepMind AlphaGo

AlphaGo marks stark difference between AI and human intelligence

Daniel Susskind

Share ▾ Author alerts ▾ Print Clip

Comments

Google's robot relied on processing power and data storage, writes Daniel Susskind



There are many ways of being smart that aren't smart like us." These are the words of Patrick Winston, a leading voice in the field of artificial intelligence. Although his idea is simple, its significance has been lost on most people thinking about the future of work. Yet this is the feature of AI that ought to preoccupy us the most.

From the 1950s to the 1980s, during the "first wave" of AI research, it was generally

What AlphaGo's Win Means for Your Job

COMMENTARY

by Howard Yu

@HowardHYu

MARCH 21, 2016, 3:11 PM EDT



And the future of artificial intelligence

Just last week, machines crossed a momentous milestone. Google's AlphaGo, a computer algorithm, beat Go world champion Lee Sedol 4 to 1 in the ancient Chinese board game.

Unlike Western chess, which consists of about 40 moves in a

AlphaGo's Success Shows the Human Advantage Is Eroding Fast



Howard Yu is a professor of strategic management and innovation at [IMD business school](#).

MARCH 9, 2016

It's impressive that computers can be programmed to conduct conversations that make them [indistinguishable](#) from a person. But Google's AlphaGo is demonstrating for the first time that machines can truly learn and think in a human way.

In 1996, IBM's Deep Blue program overwhelmed the world's greatest player, Garry Kasparov, with a brute force approach, in which the machine could account for all the possible outcomes.

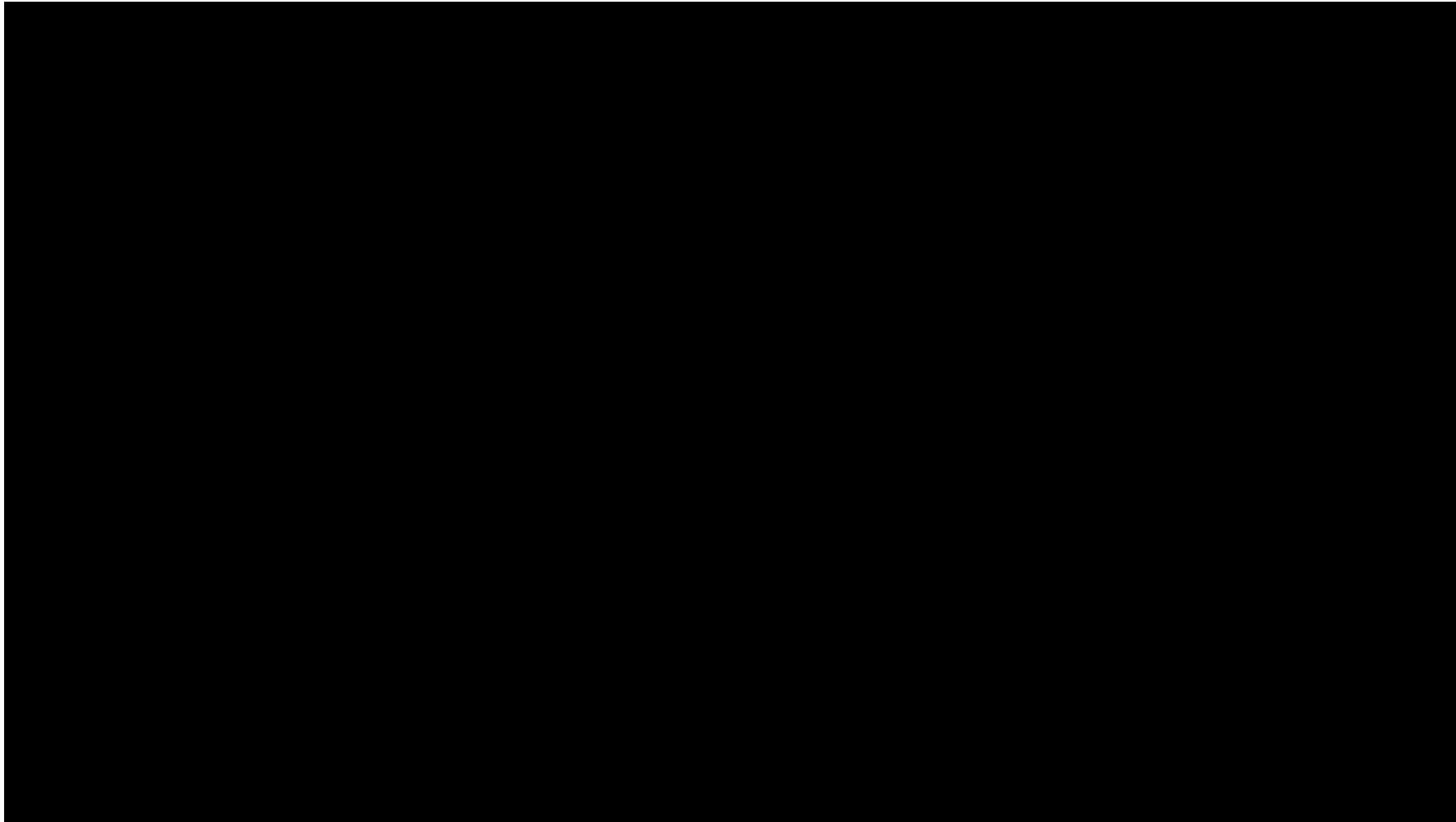
AlphaGo's approach, on the other hand, is, potentially, a game changer. It can master games by adjusting as it goes.

Unlike chess, where each move affords about 40 options, the ancient board game Go, has up to [200 choices](#). The permutation of outcomes quickly compounds to a [bewildering range](#) of choices — more than the total number of atoms in the entire observable universe.

South Korean professional Go player Lee Se-Dol reviews the match after the fourth match against Google's artificial intelligence program, AlphaGo, during the Google DeepMind Challenge Match on March 13, 2016 in Seoul, South Korea. Photograph by Google via Getty Images

RECOMMENDED FOR YOU

Baidu eye



<https://www.youtube.com/watch?v=Xe5RcJ1JY3c>

Acknowledgement



Andrew Ng: Deep Learning,
Self-Taught Learning and
Unsupervised Feature
Learning [[Youtube](#)]



Yann LeCun: Deep
Learning Tutorial, ICML,
Atlanta, 2013 [[PDF](#)]



Geoff Hinton, Yoshua
Bengio & Yann LeCun:
Deep Learning: NIPS2015
Tutorial [[PDF](#)]



Yoshua Bengio: Theano: A Python
framework for fast computation of
mathematical expressions. [[URL](#)]



Andrej Karpathy: Visualizing and
Understanding Recurrent Networks, ICLR
2016, [[PDF](#)]

Outline

- A brief history of machine learning
- Understanding the human brain
- Neural Network: Concept, implementation and challenges
- Deep Belief Network (DBN): Concept and Application
- Convolutional Neural Network (CNN): Concept and Application
- Recurrent Neural Network (RNN): Concept and Application
- Deep Learning: Strengths, weaknesses and applications
- Deep Learning: Platforms, frameworks and libraries
- Demo

Introduction

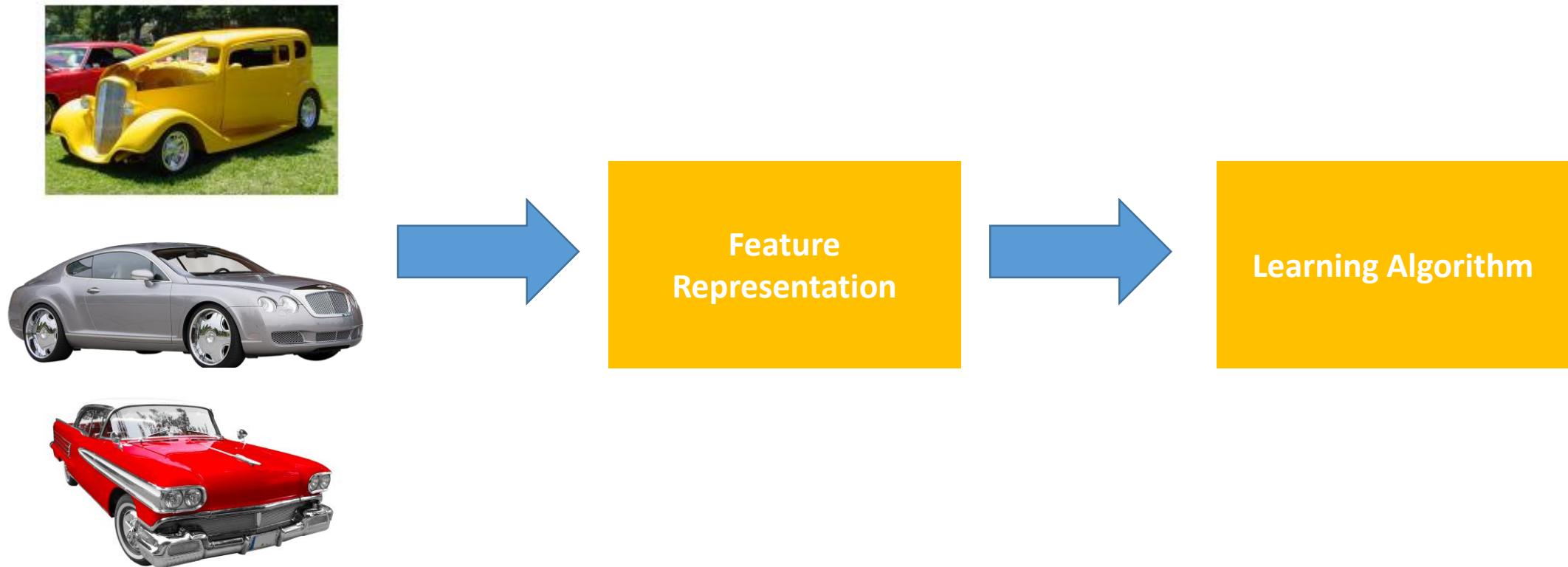
- In the past 10 years, machine learning and Artificial Intelligence have shown tremendous progress
- The recent success can be attributed to:
 - Explosion of data
 - Cheap computing cost – CPUs and GPUs
 - Improvement of machine learning models
- Much of the current excitement concerns a subfield of it called “deep learning”.



A brief history of Machine learning

- Most of the machine learning methods are based on supervised learning

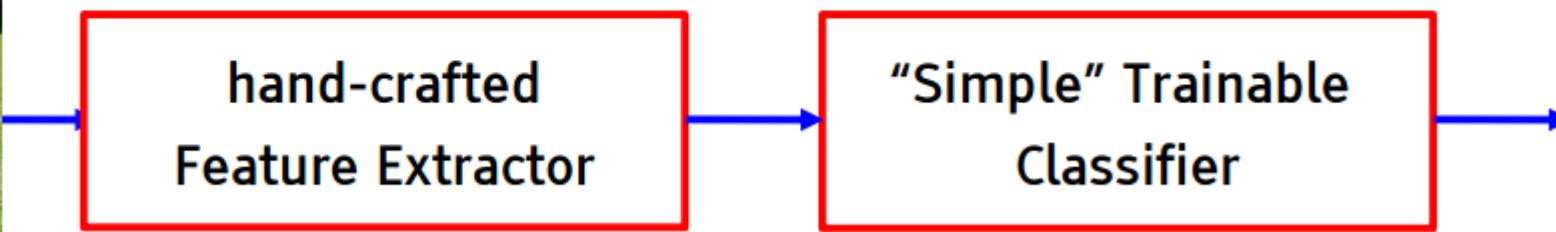
Input



A brief history of Machine learning

■ The traditional model of pattern recognition (since the late 50's)

- ▶ Fixed/engineered features (or fixed kernel) + trainable classifier



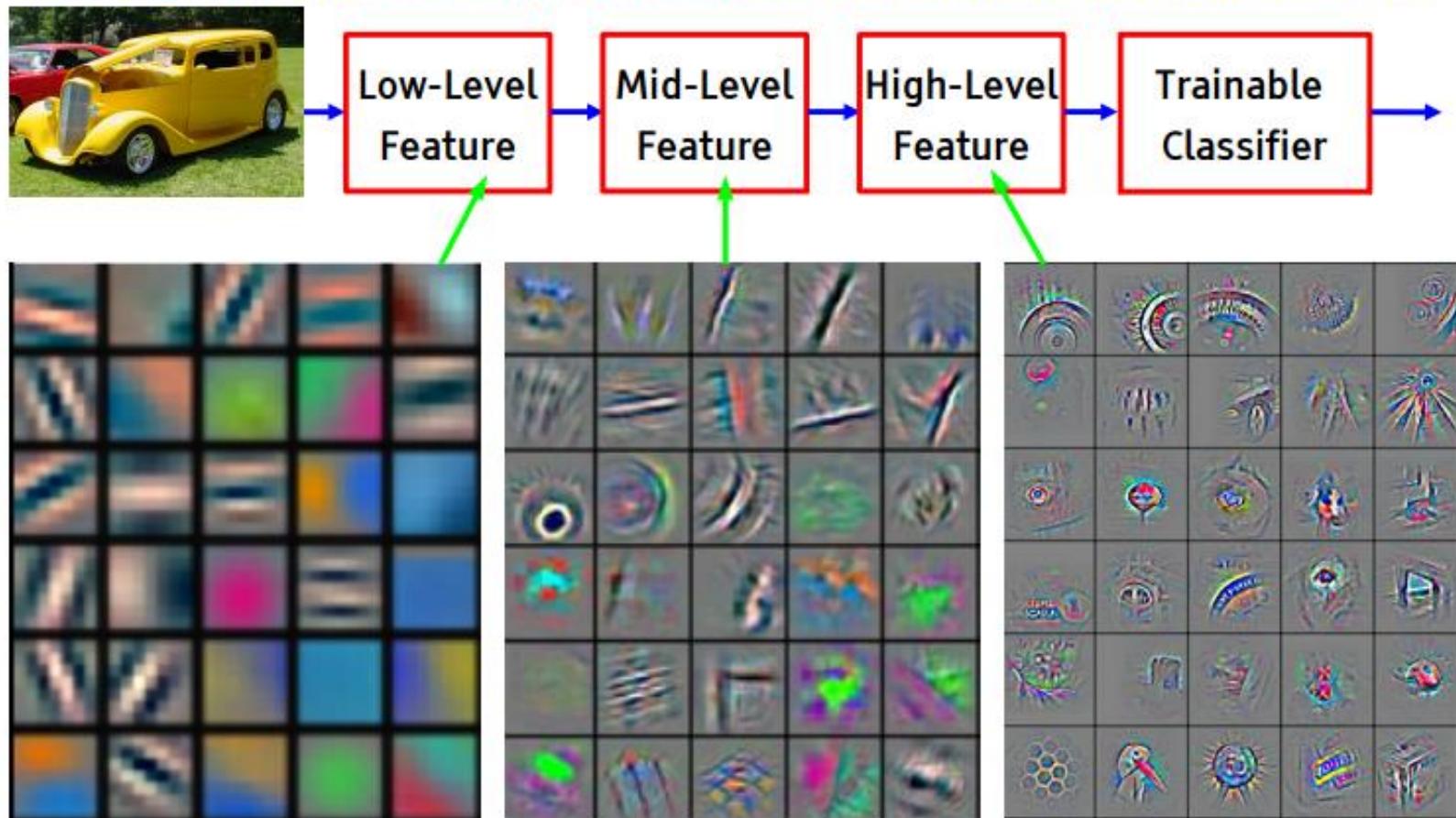
$$\begin{bmatrix} 32 & 45 & 21 & \dots \\ 12 & 10 & 45 & \dots \\ 17 & 33 & 36 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$



$$\begin{bmatrix} 12 & 56 & 18 & \dots \\ 92 & 76 & 22 & \dots \\ 33 & 63 & 71 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

Features Training

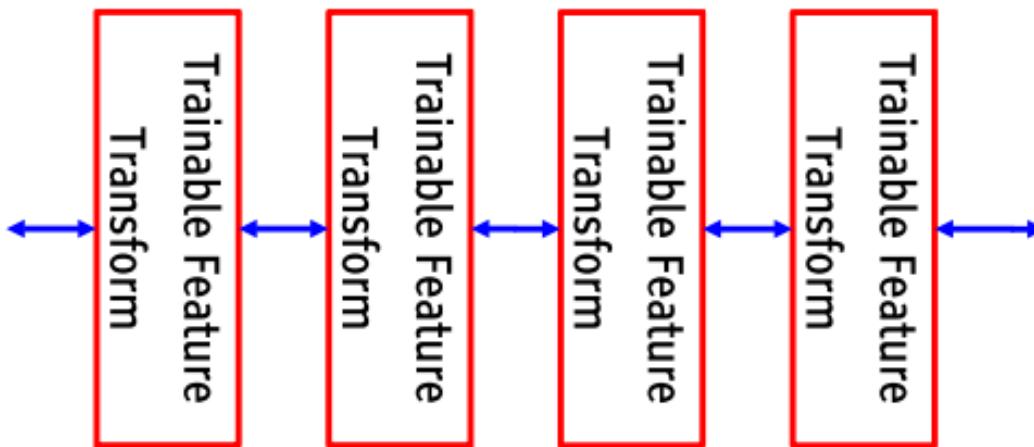
■ It's deep if it has more than one stage of non-linear feature transformation



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

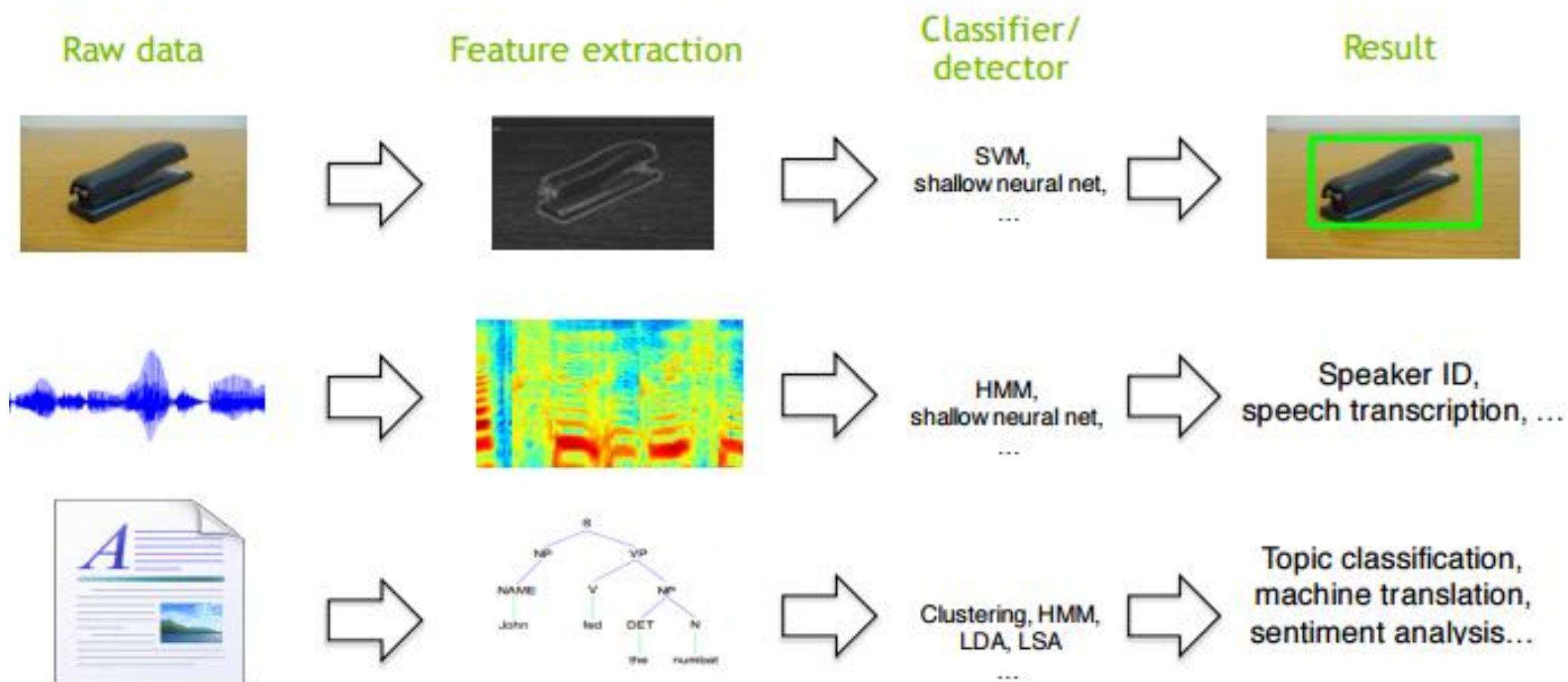
Trainable Feature Hierarchy

- Hierarchy of representations with increasing level of abstraction
- Each stage is a kind of trainable feature transform
- Image recognition
 - ▶ Pixel → edge → texton → motif → part → object
- Text
 - ▶ Character → word → word group → clause → sentence → story
- Speech
 - ▶ Sample → spectral band → sound → ... → phone → phoneme → word

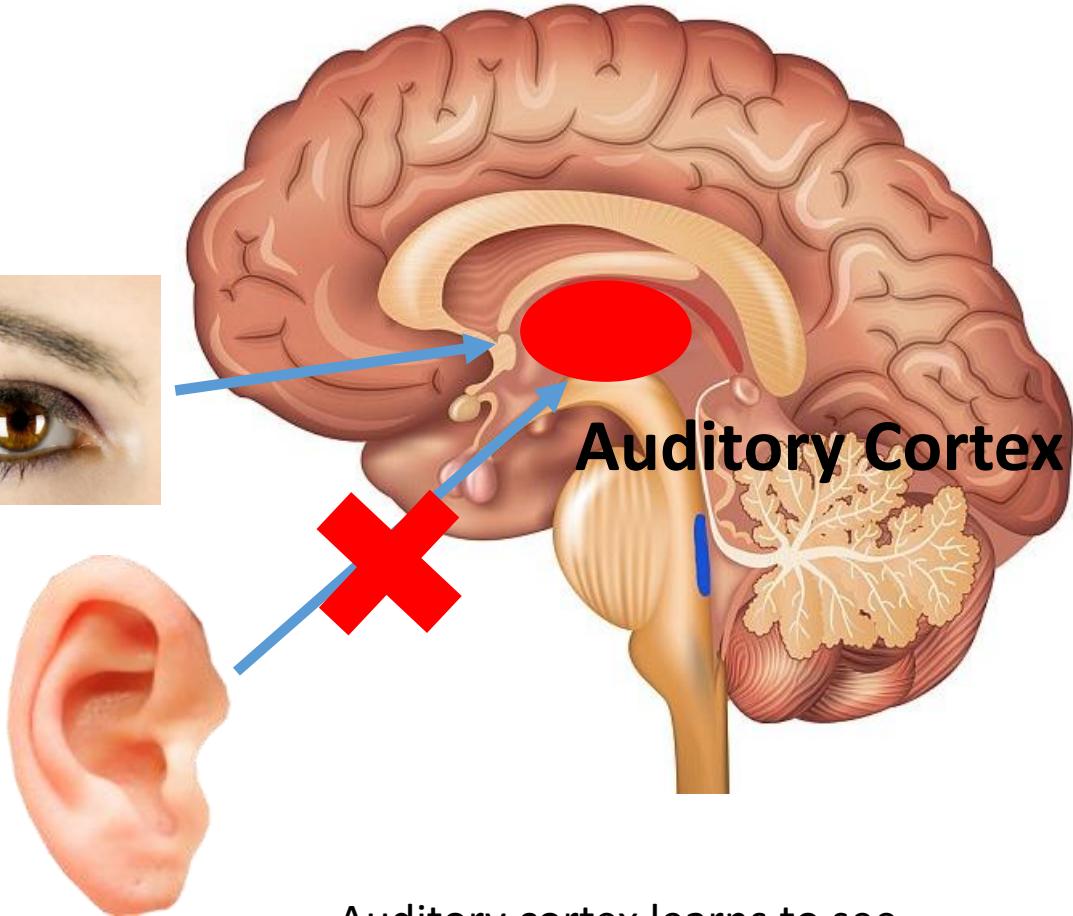


Traditional machine perception

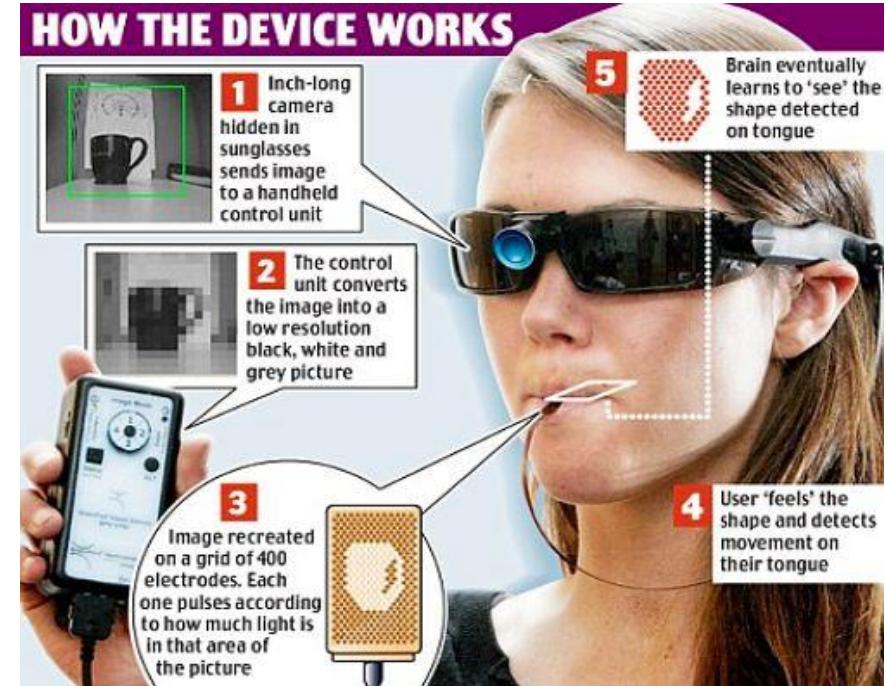
- Hand crafted feature extractors



Human Brain

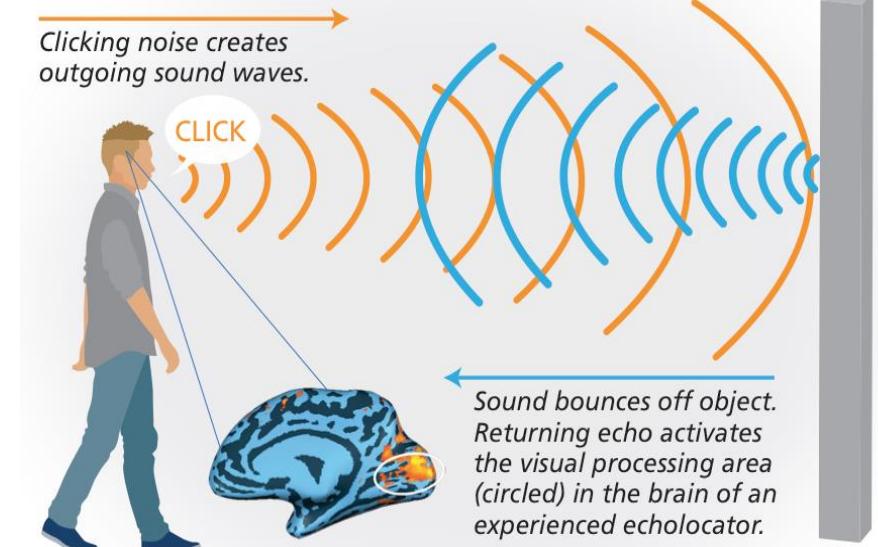


Auditory cortex learns to see.
(Same rewiring process also
works for touch/ somatosensory
cortex.)

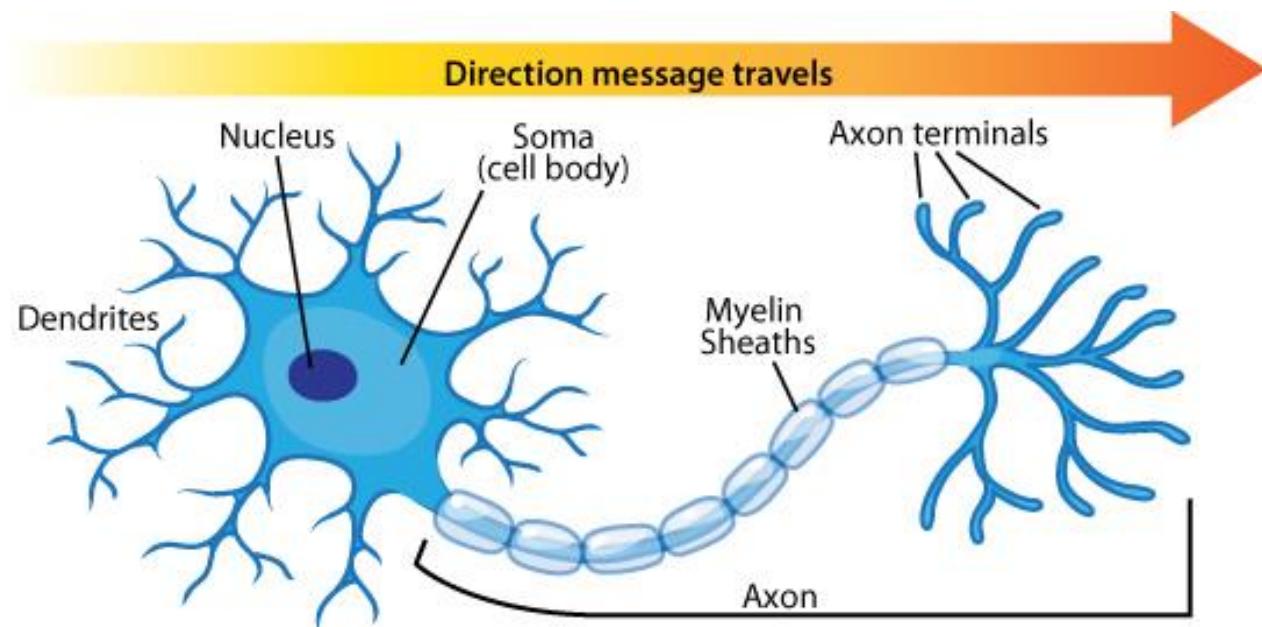


Seeing with tongue

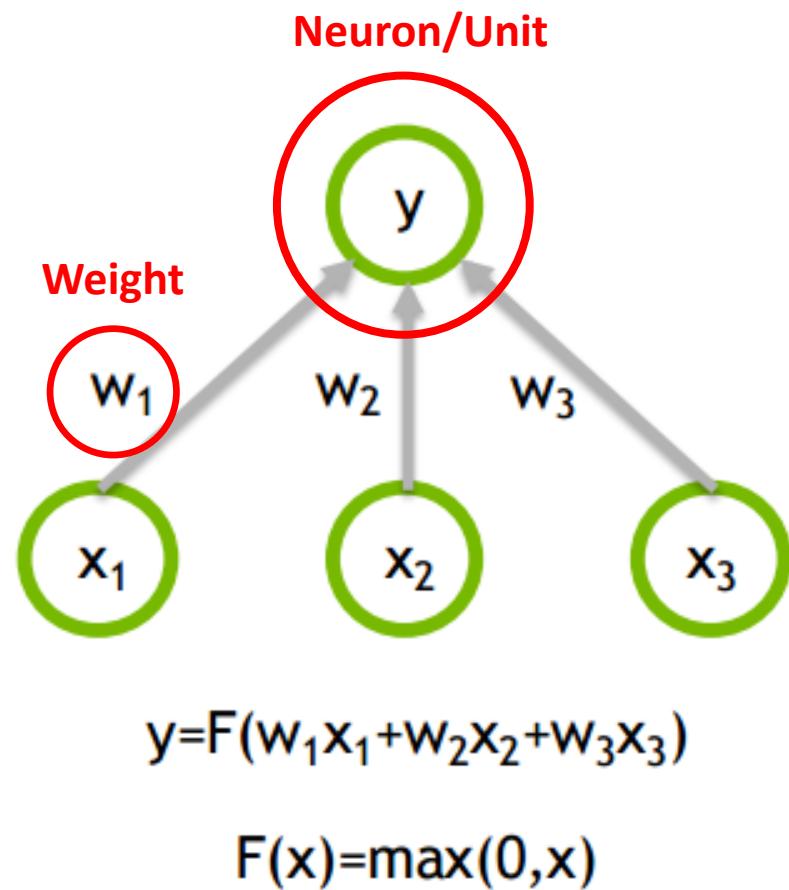
HUMAN ECHolocation: HOW IT WORKS



Human Brain



Biological Neuron

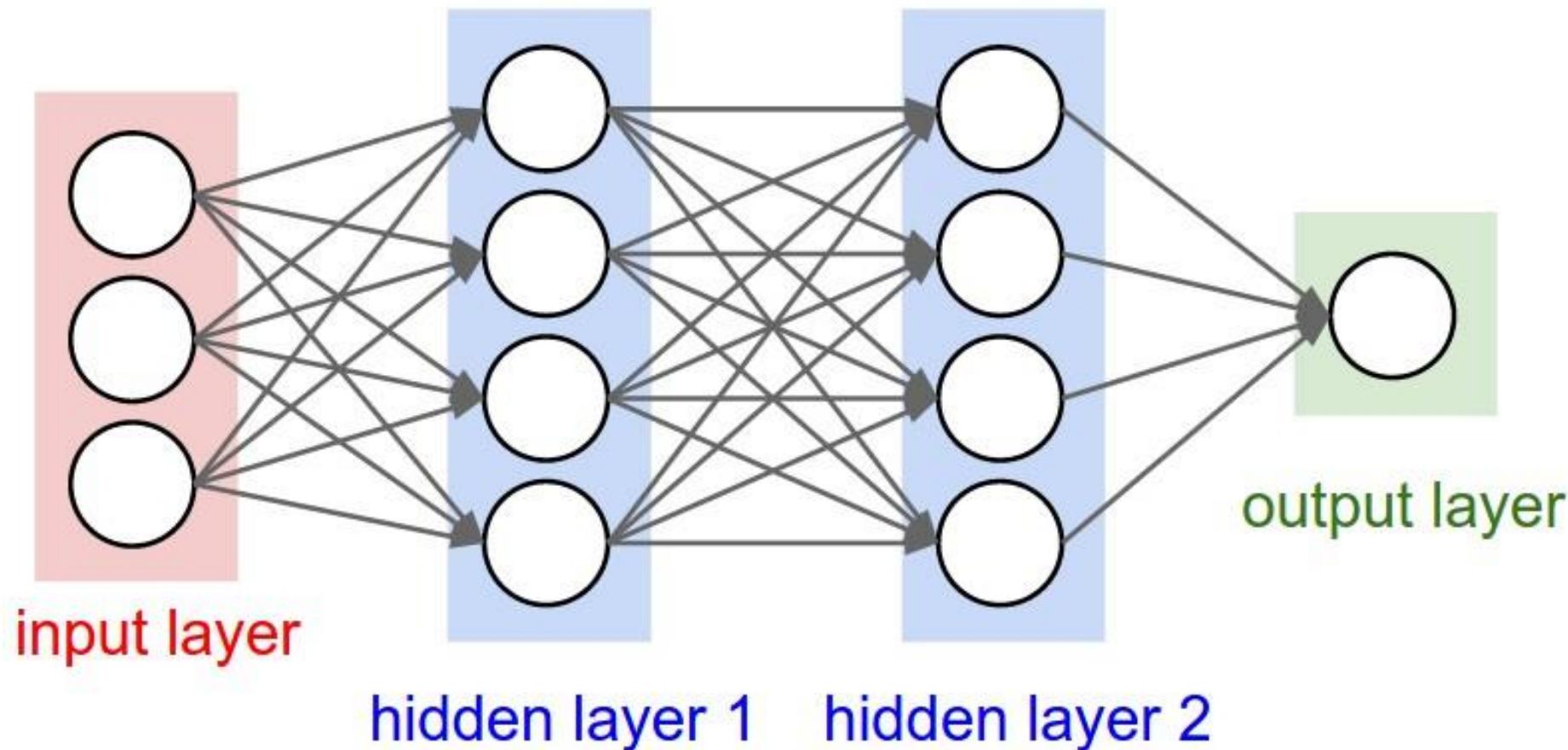


Artificial Neuron

Neural Network

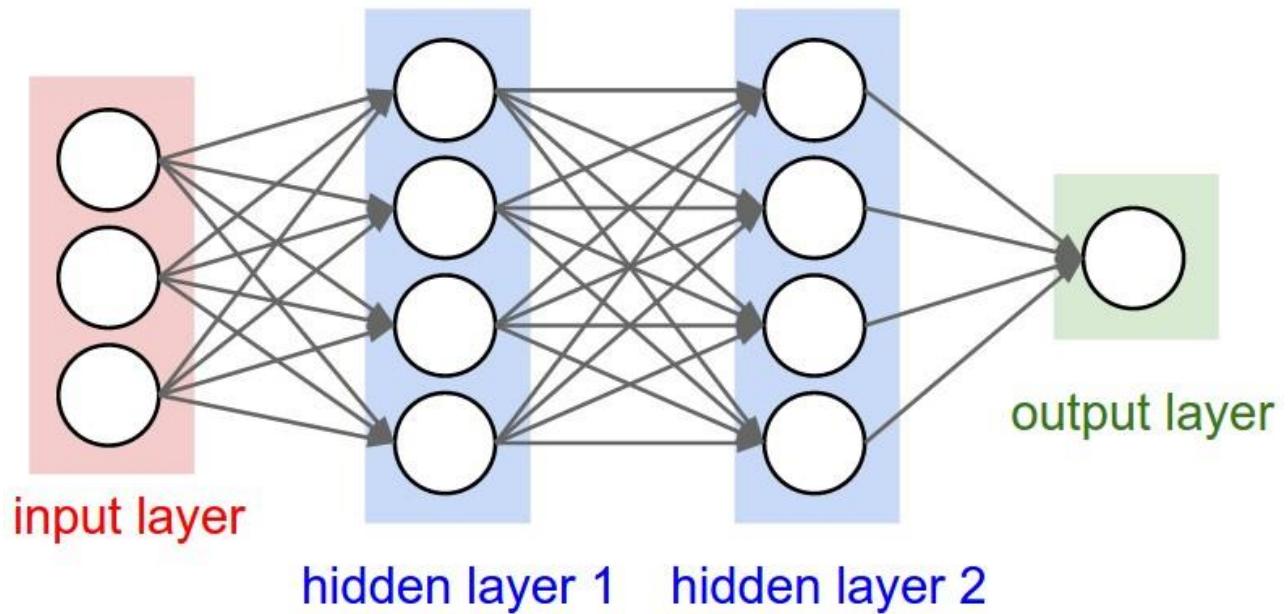
- Deep Learning is primarily about neural networks, where a network is an interconnected web of nodes and edges.
- Neural nets were designed to perform complex tasks, such as the task of placing objects into categories based on a few attributes.
- Neural nets are highly structured networks, and have three kinds of layers - an **input**, an **output**, and so called **hidden layers**, which refer to any layers between the input and the output layers.
- Each node (also called a **neuron**) in the hidden and output layers has a classifier.

Neural Network



Neural Network: Forward Propagation

- The input neurons first receive the data features of the object. After processing the data, they send their output to the first hidden layer.
- The hidden layer processes this output and sends the results to the next hidden layer.
- This continues until the data reaches the final output layer, where the output value determines the object's classification.
- This entire process is known as **Forward Propagation**, or **Forward prop**.



Neural Network: Backward Propagation

- To train a neural network over a large set of labelled data, you must continuously compute the difference between the network's predicted output and the actual output.
- This difference is called the cost, and the process for training a net is known as **backpropagation**, or **backprop**
- During backprop, **weights and biases are tweaked** slightly until the lowest possible cost is achieved.
- An important aspect of this process is the gradient, which is a measure of how much the cost changes with respect to a change in a weight or bias value.

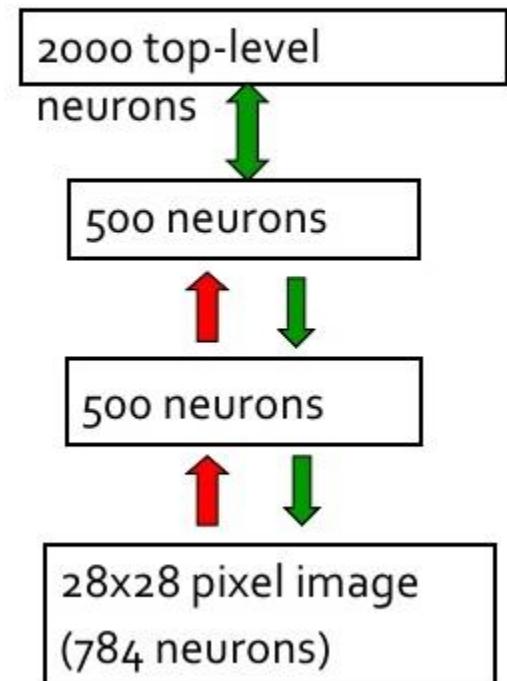
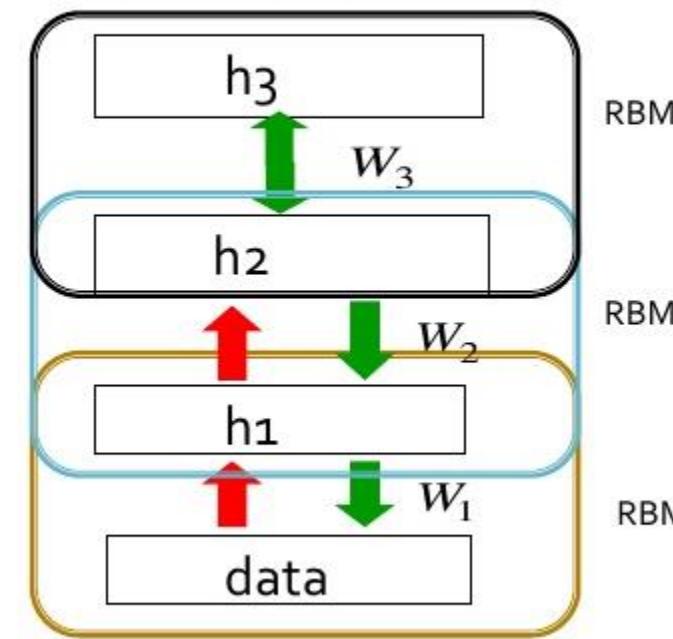
The 1990s view of what was wrong with back-propagation

- It required a lot of labelled training data
 - almost all data is unlabeled
- The learning time did not scale well
 - It was very slow in networks with multiple hidden layers.
- It got stuck at local optima
 - These were often surprisingly good but there was no good theory

Deep Belief Network (DBN)

- The Deep Belief Network, or DBN, was also conceived by Geoff Hinton.
- Used by Google for their work on the image recognition problem.
- DBN is trained two layers at a time, and these two layers are treated like an RBM.
- Throughout the net, the hidden layer of an RBM acts as the input layer of the adjacent one. So the first RBM is trained, and its outputs are then used as inputs to the next RBM. This procedure is repeated until the output layer is reached.

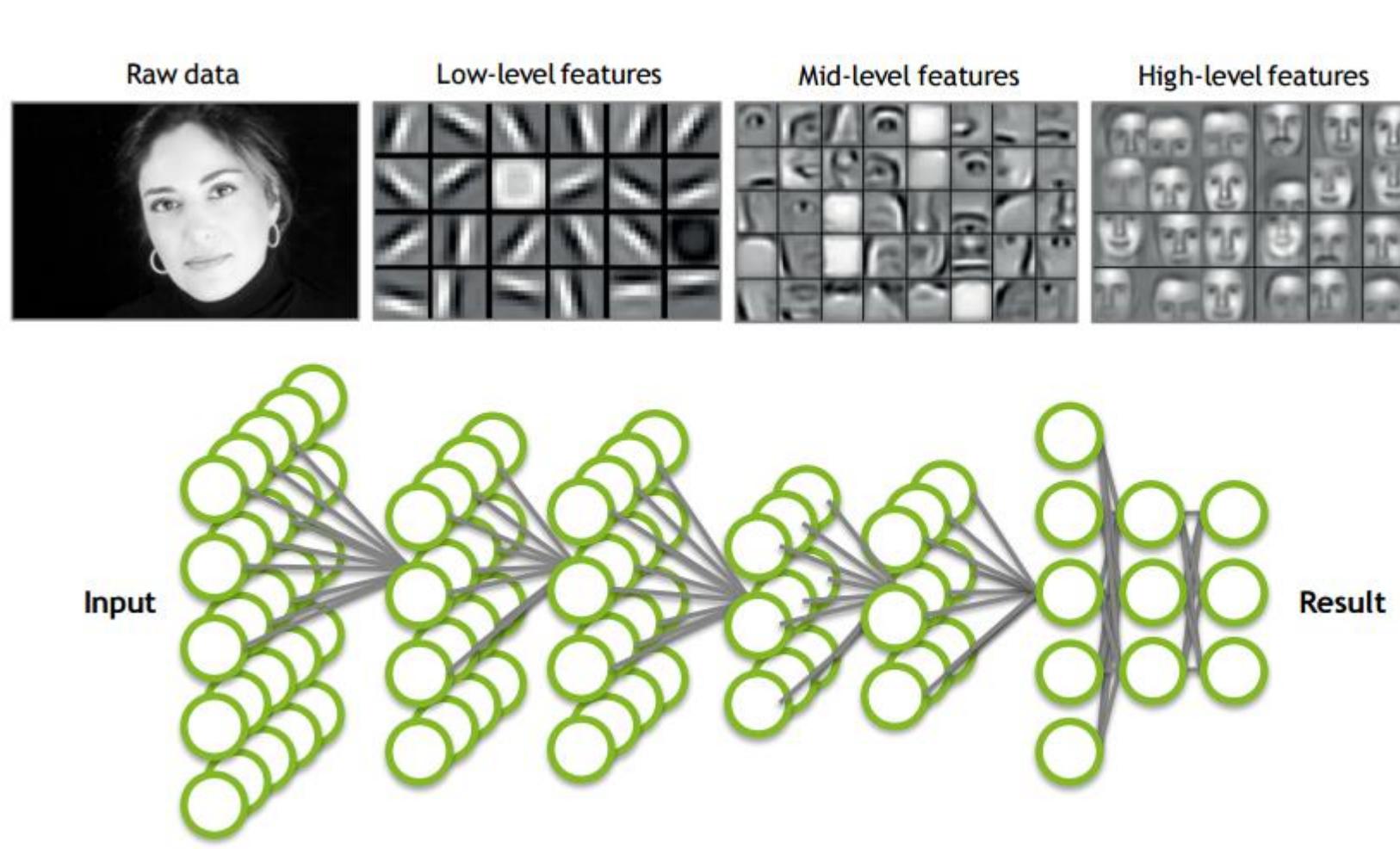
DBNs are stacks of restricted Boltzmann machines forming deep (multi-layer) architecture.



Deep Belief Network (DBN)

- DBN is capable of recognizing the inherent patterns in the data. In other words, it's a sophisticated, multilayer feature extractor.
- The unique aspect of this type of net is that each layer ends up learning the full input structure.
- Layers generally learn progressively complex patterns – for facial recognition, early layers could detect edges and later layers would combine them to form facial features.
- DBN learns the hidden patterns globally, like a camera slowly bringing an image into focus.
- DBN still requires a set of labels to apply to the resulting patterns. As a final step, the DBN is fine-tuned with supervised learning and a small set of labeled examples.

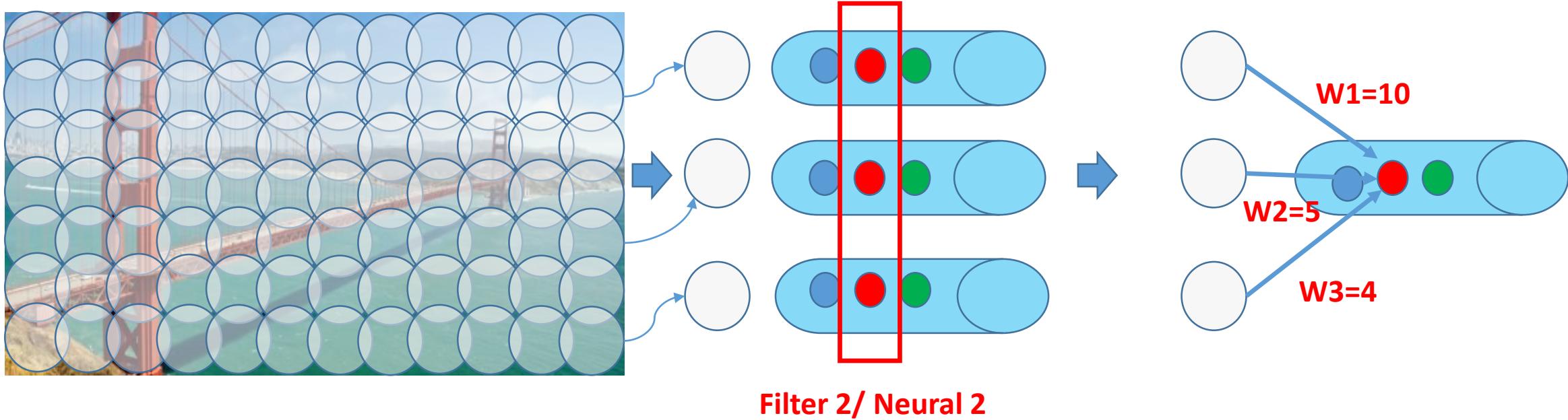
Deep Neural Network (Deep Net)



Convolutional Neural Network (CNN)

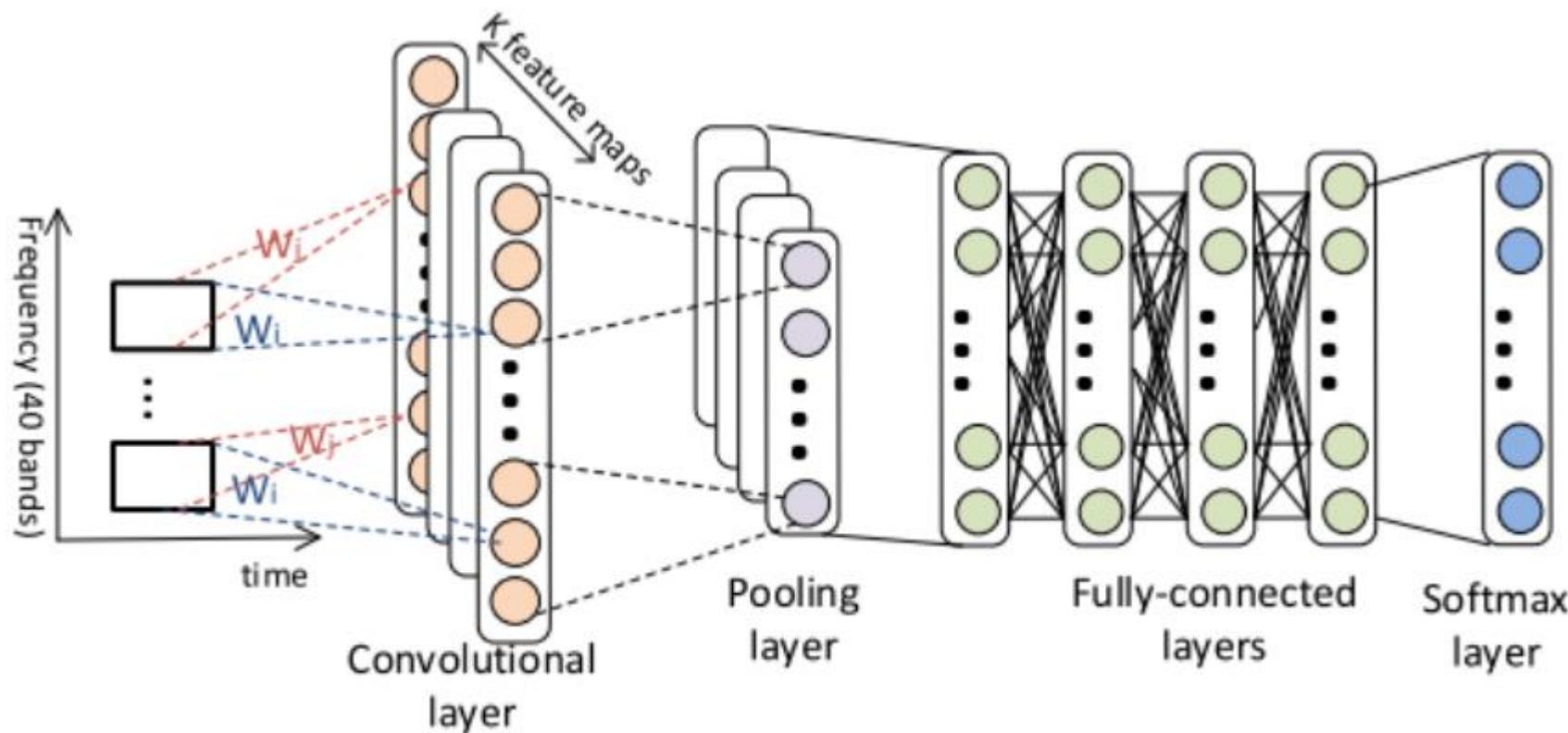
- CNN inspired by the Visual Cortex.
- CNNs are deep nets that are used for image, object, and even speech recognition.
- Pioneered by Yann Lecun (NYU)
- Deep supervised neural networks are generally too difficult to train.
- CNNs have multiple types of layers, the first of which is the convolutional layer.

Convolutional Neural Network (CNN)



- A series of filters forms layer one, called the convolutional layer. The weights and biases in this layer determine the effectiveness of the filtering process.
- Each flashlight represents a single neuron. Typically, neurons in a layer activate or fire. On the other hand, in the convolutional layer, neurons search for patterns through convolution. Neurons from different filters search for different patterns, and thus they will process the input differently.

Convolutional Neural Network (CNN)

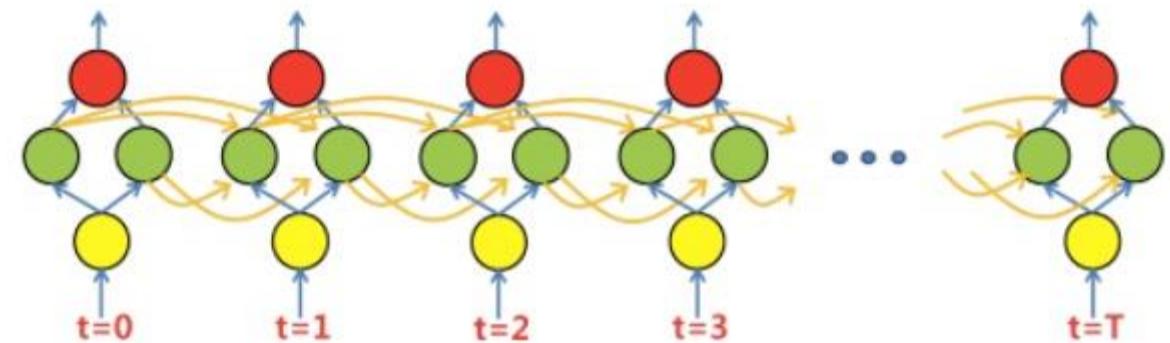


CNN: Application

- Classify a scene in an image
 - Image Classifier Demo (NYU): <http://horatio.cs.nyu.edu/>
- Describe or understanding an image
 - Toronto Deep Learning Demo: <http://deeplearning.cs.toronto.edu/i2t>
 - MIT Scene Recognition Demo: <http://places.csail.mit.edu/demo.html>
- Handwriting recognition
 - Handwritten digits recognition:
<http://cs.stanford.edu/people/karpathy/convnetjs/demo/mnist.html>
- Video classification
 - Large-scale Video Classification with Convolutional Neural Networks
<http://cs.stanford.edu/people/karpathy/deepvideo/>

Recurrent Neural Network (RNN)

- The Recurrent Neural Net (RNN) is the brainchild of Juergen Schmidhuber and Sepp Hochreiter.
- RNNs have a feedback loop where the net's output is fed back into the net along with the next input.
- RNNs receive an input and produce an output. Unlike other nets, the inputs and outputs can come in a sequence.
- Variant of RNN is Long Term Short Memory (LSTM)



RNN: Application

- RNN is suitable for time series data, where an output can be the next value in a sequence, or the next several values

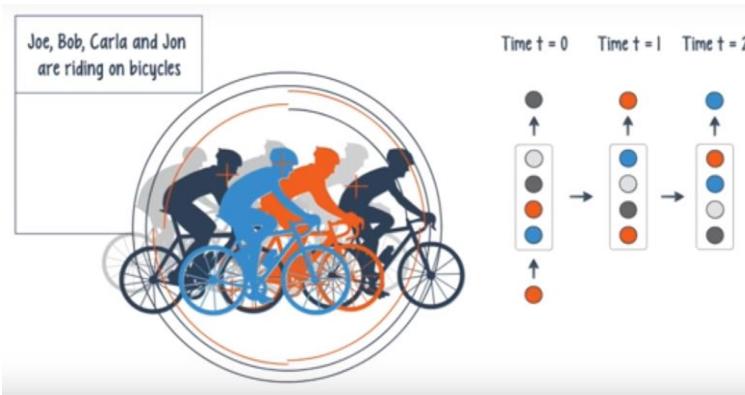
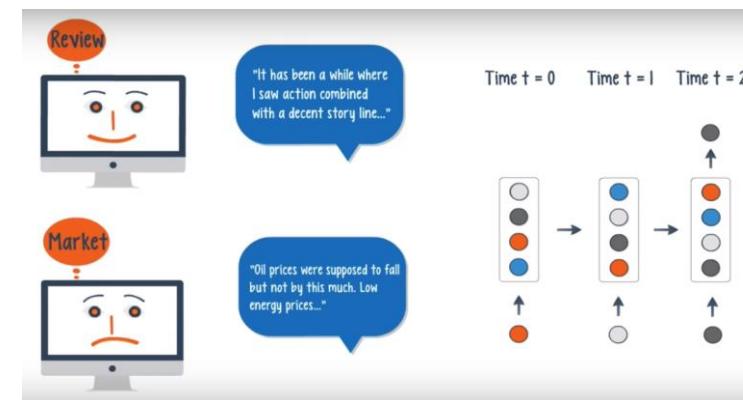
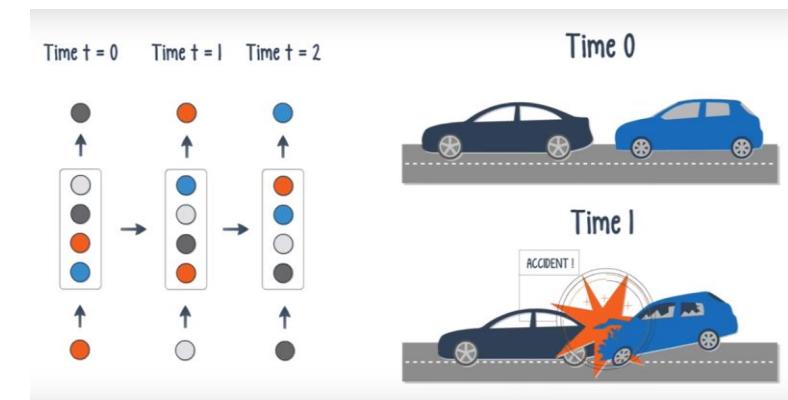


Image captioning



Document Classification



Classify Image frame by frame

Deep Learning: Benefits

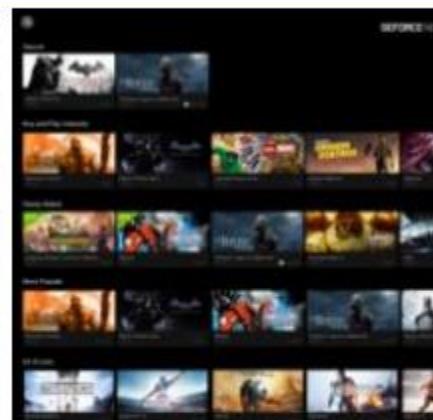
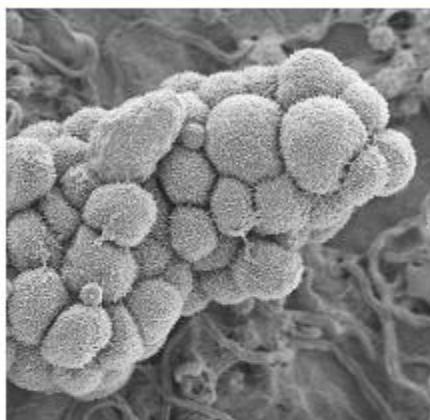
- Robust
 - No need to design the features ahead of time – features are automatically learned to be optimal for the task at hand
 - Robustness to natural variations in the data is automatically learned
- Generalizable
 - The same neural net approach can be used for many different applications and data types
- Scalable
 - Performance improves with more data, method is massively parallelizable

Deep Learning: Weaknesses

- Deep Learning **requires a large dataset**, hence long training period.
- In term of cost, Machine Learning methods like SVMs and other tree ensembles are very easily deployed even by relative machine learning novices and can usually get you reasonably good results.
- Deep learning methods **tend to learn everything**. It's better to encode prior knowledge about structure of images (or audio or text).
- The learned features are often **difficult to understand**. Many vision features are also not really human-understandable (e.g, concatenations/combinations of different features).
- Requires **a good understanding of how to model** multiple modalities with traditional tools.

Deep Learning: Applications

DEEP LEARNING EVERYWHERE



INTERNET & CLOUD

Image Classification
Speech Recognition
Language Translation
Language Processing
Sentiment Analysis
Recommendation

MEDICINE & BIOLOGY

Cancer Cell Detection
Diabetic Grading
Drug Discovery

MEDIA & ENTERTAINMENT

Video Captioning
Video Search
Real Time Translation

SECURITY & DEFENSE

Face Detection
Video Surveillance
Satellite Imagery

AUTONOMOUS MACHINES

Pedestrian Detection
Lane Tracking
Recognize Traffic Sign

Deep Learning: Application



<https://deepmind.com/alpha-go>



[Robotic grasping](#)

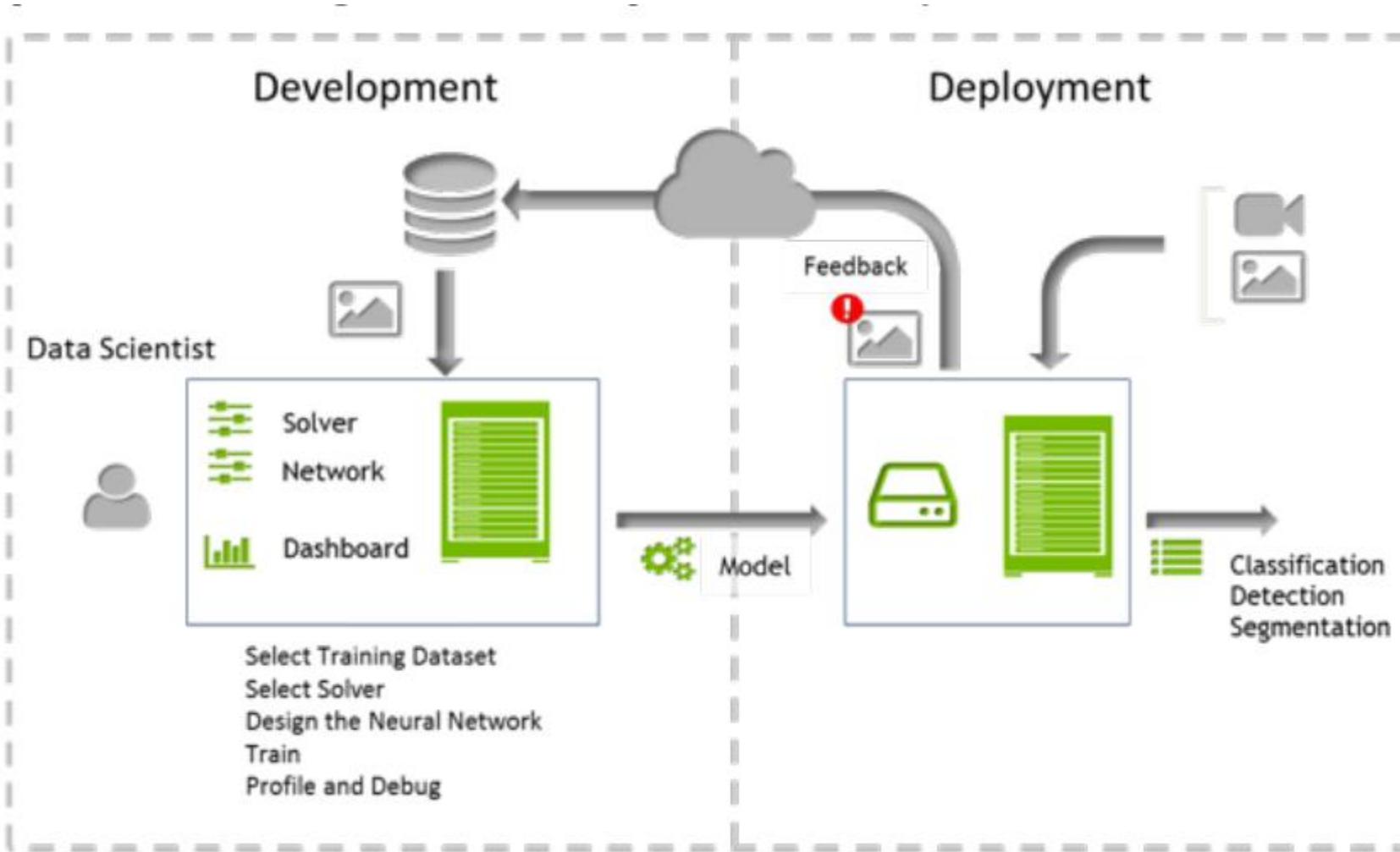
A screenshot of the "MIT Scene Recognition Demo" interface. The title "MIT Scene Recognition Demo" is at the top. Below it is a descriptive paragraph about the demo's purpose and how it uses Places-CNN. There are three input fields: "Upload:" with a "Choose File" button, "URL:" with a text input field containing "http://places.csail.mit.edu/demo/19" and a "Run" button, and "Click One:" with a grid of thumbnail images representing different scenes. A heatmap overlay is visible on the right side of the interface.

<http://places.csail.mit.edu/demo.html>



[Pedestrian detection using DL](#)

Deep Learning Development Cycle



Deep Learning: Platform & Frameworks & Libraries

Platform

- Ersatz Labs - cloud-based deep learning platform [<http://www.ersatz1.com/>]
- H2O – deep learning framework that comes with R and Python interfaces [<http://www.h2o.ai/verticals/algos/deep-learning/>]

Framework

- Caffe - deep learning framework made with expression, speed, and modularity in mind. Developed by the Berkeley Vision and Learning Center (BVLC) [<http://caffe.berkeleyvision.org/>]
- Torch - scientific computing framework with wide support for machine learning algorithms that puts GPUs first. Based on Lua programming language [<http://torch.ch/>]

Library

- Tensorflow - open source software library for numerical computation using data flow graphs from Google [<https://www.tensorflow.org/>]
- Theano - a python library developed by Yoshua Bengio's team [<http://deeplearning.net/software/theano/>]

Learned Models

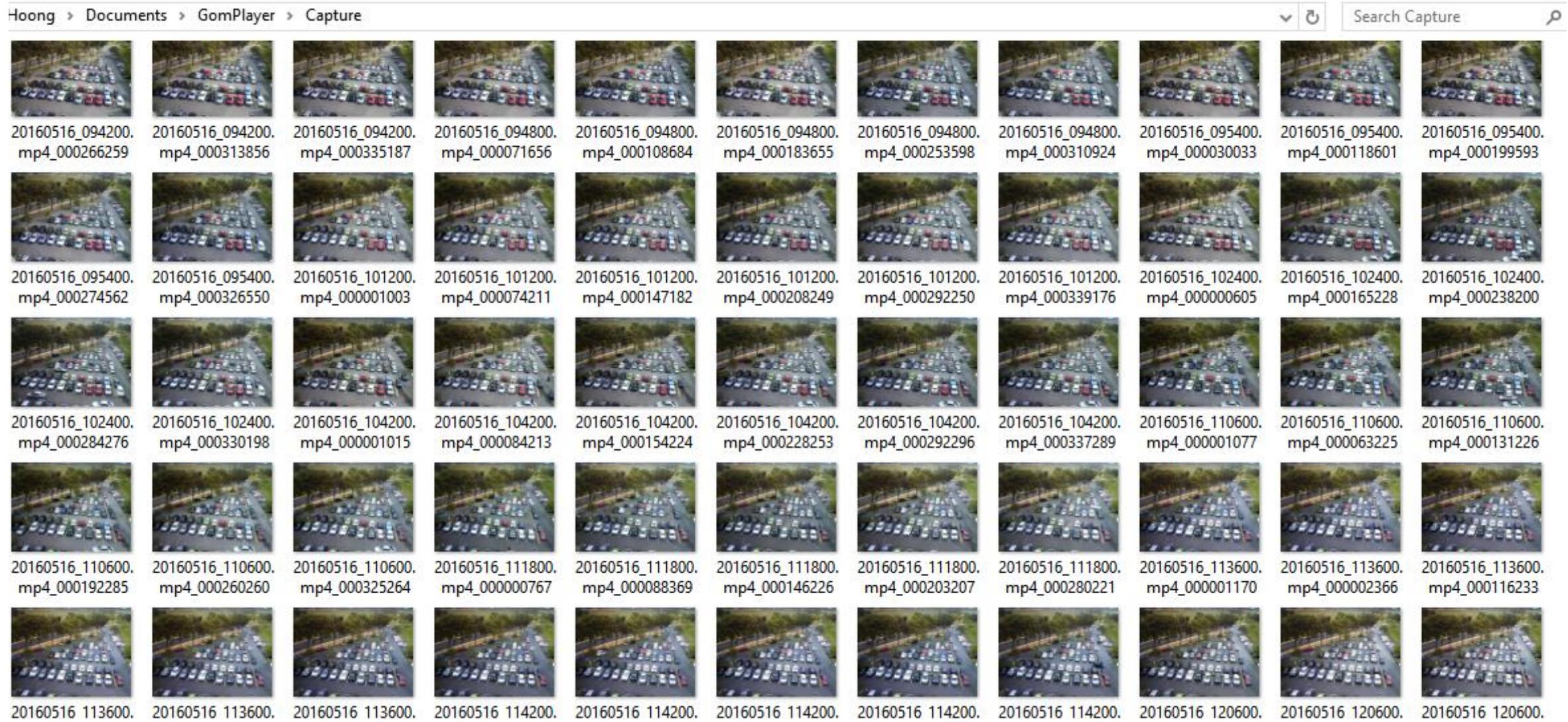
- Trained Models can be shared with others
- Save the training time
- For example: AlexNet, GoogLeNet, ParseNet, etc
- URLs:
 - <https://github.com/BVLC/caffe/wiki/Model-Zoo>
 - <http://deeplearning4j.org/model-zoo>

DEMO

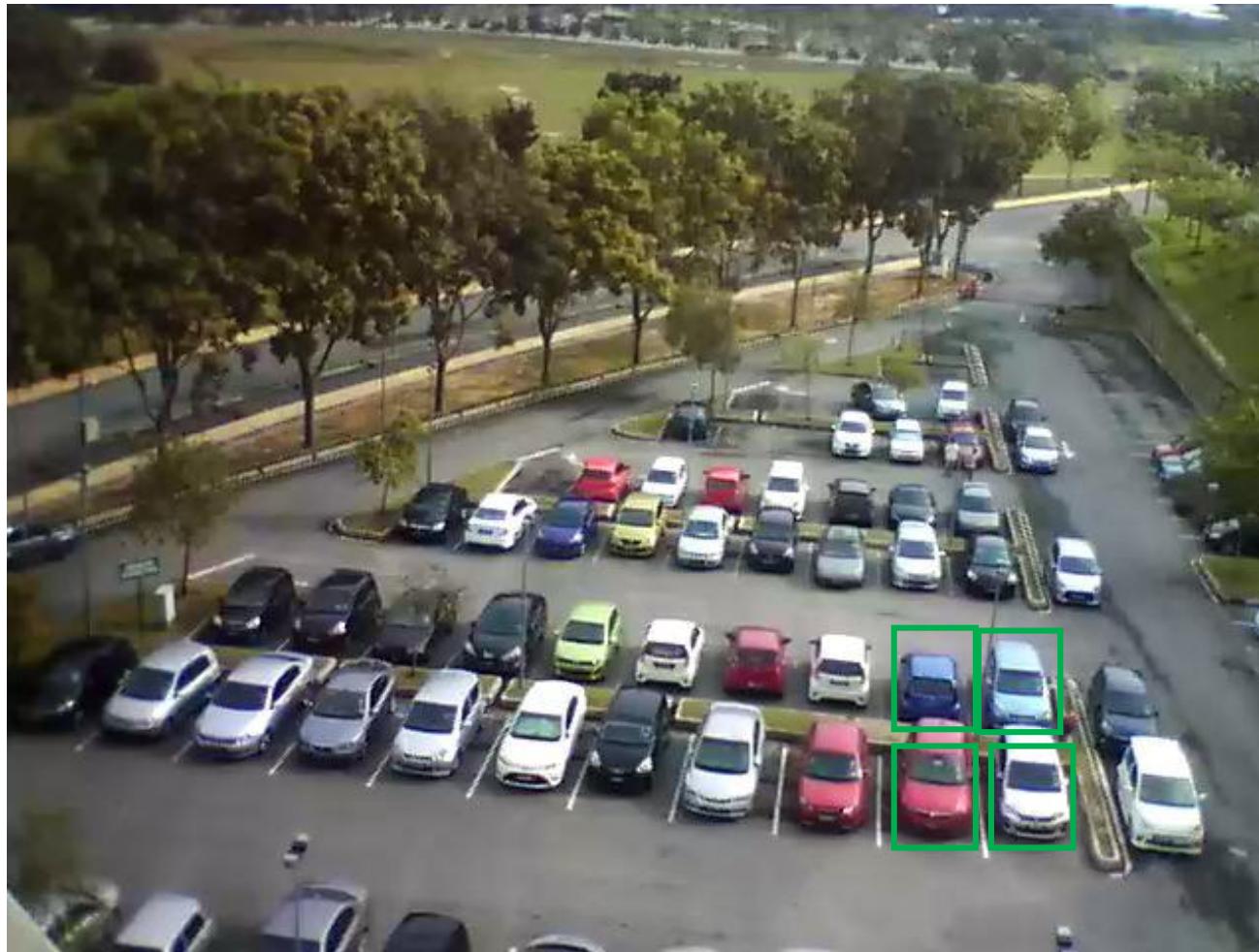
Nvidia: Digits

- The NVIDIA Deep Learning GPU Training System (DIGITS) puts the power of deep learning in the hands of data scientists and researchers.
- Quickly design the best deep neural network (DNN) for your data using real-time network behavior visualization.
- <https://developer.nvidia.com/digits>

Car Park Images



Car park images



Cropped Car Park space

jan Hoong > Pictures > carpark > cp02 >

Search cp02

20160516_094200.
mp4_000266259

20160516_094200.
mp4_000313856

20160516_094200.
mp4_000335187

20160516_094800.
mp4_000071656

20160516_094800.
mp4_000108684

20160516_094800.
mp4_000183655

20160516_094800.
mp4_000253598

20160516_094800.
mp4_000310924

20160516_095400.
mp4_000030033

20160516_095400.
mp4_000118601

20160516_095400.
mp4_000199593

jan Hoong > Pictures > carpark > cp04 >

Search cp04

20160516_094200.
mp4_000266259

20160516_094200.
mp4_000313856

20160516_094200.
mp4_000335187

20160516_094800.
mp4_000071656

20160516_094800.
mp4_000108684

20160516_094800.
mp4_000183655

20160516_094800.
mp4_000253598

20160516_094800.
mp4_000310924

20160516_094800.
mp4_000030033

20160516_095400.
mp4_000118601

20160516_095400.
mp4_000199593

20160516_095400.
mp4_000274562

20160516_095400.
mp4_000326550

20160516_101200.
mp4_000001003

20160516_101200.
mp4_00007421

20160516_095400.
mp4_000274562

20160516_095400.
mp4_000326550

20160516_101200.
mp4_000001003

20160516_101200.
mp4_000074211

20160516_101200.
mp4_000147182

20160516_101200.
mp4_000208249

20160516_101200.
mp4_000292250

20160516_101200.
mp4_000339176

20160516_102400.
mp4_000000605

20160516_102400.
mp4_000165228

20160516_102400.
mp4_000238200

20160516_102400.
mp4_000284276

20160516_102400.
mp4_000330198

20160516_104200.
mp4_000001015

20160516_104200.
mp4_00008421

20160516_102400.
mp4_000284276

20160516_102400.
mp4_000330198

20160516_104200.
mp4_000001015

20160516_104200.
mp4_000084213

20160516_104200.
mp4_000154224

20160516_104200.
mp4_000228253

20160516_104200.
mp4_000292296

20160516_104200.
mp4_000337289

20160516_110600.
mp4_000001077

20160516_110600.
mp4_000063225

20160516_110600.
mp4_000131226

20160516_110600.
mp4_000192285

20160516_110600.
mp4_000260260

20160516_110600.
mp4_000325264

20160516_111800.
mp4_0000000767

20160516_111800.
mp4_000088369

20160516_111800.
mp4_000146226

20160516_111800.
mp4_000203207

20160516_111800.
mp4_000280221

20160516_113600.
mp4_000001170

20160516_113600.
mp4_000002366

20160516_113600.
mp4_000116233

20160516_113600.
mp4_000206211

20160516_113600.
mp4_000291169

20160516_113600.
mp4_000335210

20160516_114200.
mp4_0000000735

20160516_114200.
mp4_000063257

20160516_114200.
mp4_000126213

20160516_114200.
mp4_000191250

20160516_114200.
mp4_000296167

20160516_120600.
mp4_000001524

20160516_120600.
mp4_000087224

20160516_120600.
mp4_000166275

20160516_120600.
mp4_000223362

20160516_120600.
mp4_000287345

20160516_121800.
mp4_000000797

20160516_121800.
mp4_000076212

20160516_121800.
mp4_000076787

20160516_121800.
mp4_000170255

20160516_121800.
mp4_000259775

20160516_121800.
mp4_000323218

20160516_124200.
mp4_000001204

20160516_124200.
mp4_000067311

Digits – Image Classification Model

New Image Classification Model - Mozilla Firefox
localhost/models/images/classification/new

DIGITS New Model Version 3.0.0

New Image Classification Model

Select Dataset ?
 carpark
 MNIST

carpark
Done 08:06:01 AM
Image Size 51x56
Image Type COLOR
DB backend lmdb
Create DB (train) 2400 images
Create DB (val) 800 images

Use client side file
Python Layer File (server side) ?

Data Transformations
Crop Size ? none
Subtract Mean ? Image

Solver Options
Training epochs ? 30
Snapshot interval (in epochs) ? 1
Validation interval (in epochs) ? 1
Random seed ? [none]

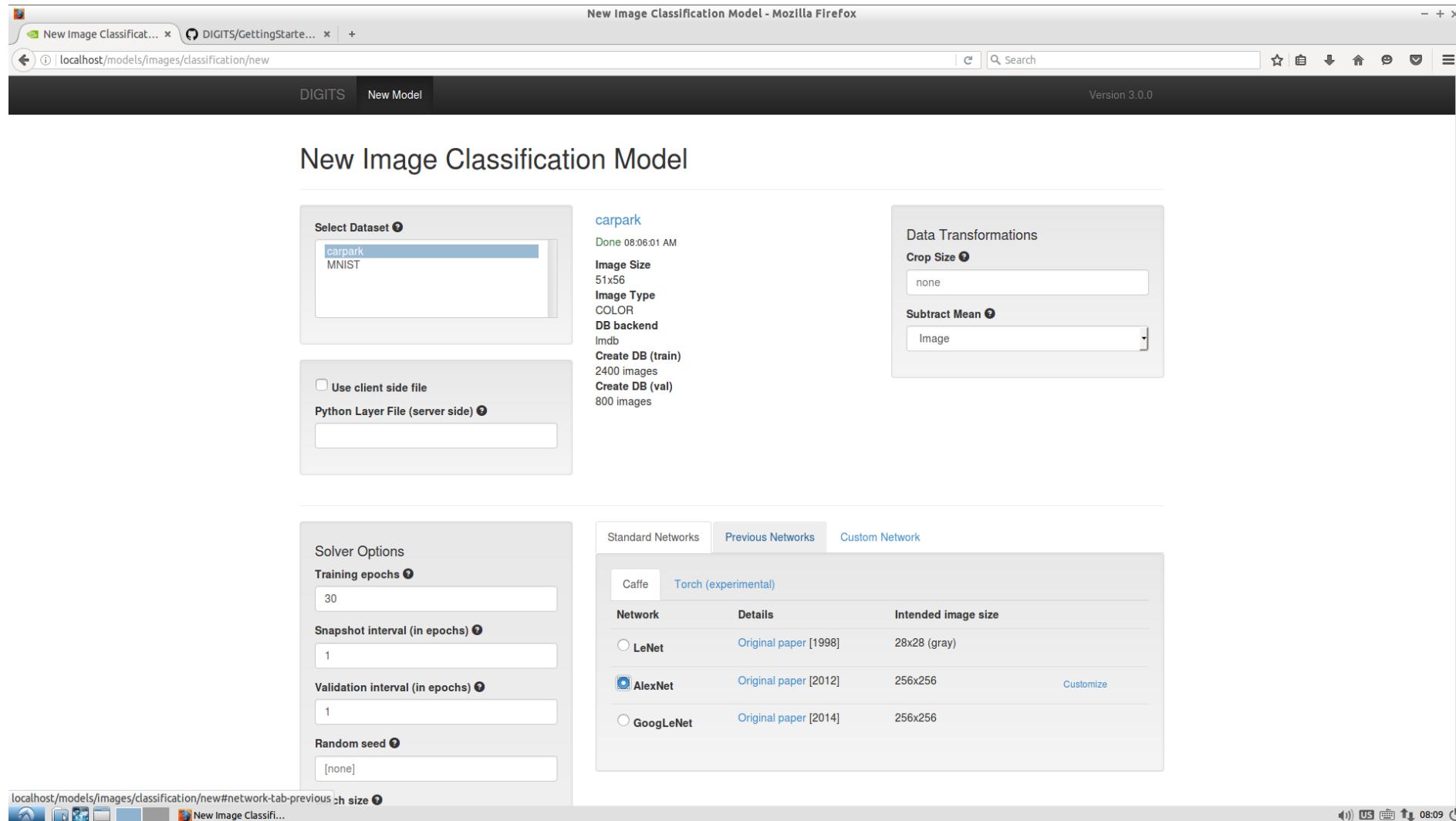
Standard Networks Previous Networks Custom Network

Caffe Torch (experimental)

Network	Details	Intended image size
<input type="radio"/> LeNet	Original paper [1998]	28x28 (gray)
<input checked="" type="radio"/> AlexNet	Original paper [2012]	256x256
<input type="radio"/> GoogLeNet	Original paper [2014]	256x256

localhost/models/images/classification/new#network-tab-previous ch size ?

08:09



Digits – AlexNet Training

AlexNet-carpark - Mozilla Firefox

AlexNet-carpark DIGITS/GettingStart... +

localhost/models/20160718-081011-9f13

Search

Version 3.0.0

DIGITS Image Classification Model

AlexNet-carpark

Image Classification Model

Clone Job Abort Job Delete Job

Disk Size
0 B

Network (train/val)
[train_val.prototxt](#)

Network (deploy)
[deploy.prototxt](#)

Solver
[solver.prototxt](#)

Raw caffe output
[caffe_output.log](#)

Dataset

carpark

Done 08:06:01 AM

Image Size
51x56

Image Type
COLOR

DB backend
lmdb

Create DB (train)
2400 images

Create DB (val)
800 images

Job Status Running

- Initialized at 08:10:11 AM (1 second)
- Running at 08:10:12 AM

Train Caffe Model Running ▾

2%

Estimated time remaining: 10 minutes, 35 seconds

- Initialized at 08:10:11 AM (1 second)
- Running at 08:10:12 AM

GPU Usage

Quadro K600 (#0)

Memory
479 MB / 1020 MB (46.8%)

GPU Utilization
99%

Temperature
59 °C

Notes

None

08:10

The figure shows a dual-axis plot of Loss and Accuracy over time. The left y-axis represents Loss from 0.1 to 0.7. The right y-axis represents Accuracy (%) from 10 to 100. The plot shows a general downward trend in loss and an upward trend in accuracy. There is a notable dip in loss around the 40-second mark and a peak in accuracy around the 70-second mark.

Time	Loss	Accuracy (%)
0	0.70	70
10	0.65	75
20	0.60	80
30	0.55	85
40	0.40	90
50	0.50	95
60	0.60	98
70	0.68	100
80	0.65	95
90	0.55	85
100	0.50	75

Digits – AlexNet Training

AlexNet-carpark - Mozilla Firefox

AlexNet-carpark DIGITS/GettingStart... localhost/models/20160718-081011-9f13

DIGITS Image Classification Model Version 3.0.0

AlexNet-carpark

Image Classification Model

Clone Job Abort Job Delete Job

Disk Size
0 B

Network (train/val)
[train_val.prototxt](#)

Network (deploy)
[deploy.prototxt](#)

Solver
[solver.prototxt](#)

Raw caffe output
[caffe_output.log](#)

Dataset
carpark

Done 08:06:01 AM

Image Size
51x56

Image Type
COLOR

DB backend
lmdb

Create DB (train)
2400 images

Create DB (val)
800 images

Job Status Running

- Initialized at 08:10:11 AM (1 second)
- Running at 08:10:12 AM

Train Caffe Model Running ▾
28%

Estimated time remaining: 4 minutes, 31 seconds

- Initialized at 08:10:11 AM (1 second)
- Running at 08:10:12 AM

GPU Usage

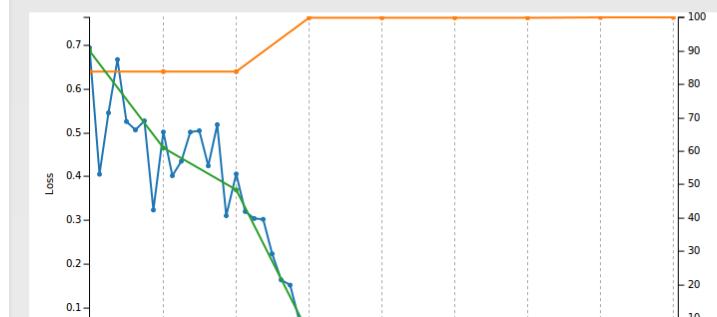
Quadro K600 (#0)

Memory
479 MB / 1020 MB (46.8%)

GPU Utilization
99%

Temperature
68 °C

Notes
None



AlexNet-carpark ... Software Updater

08:11

Digits – Testing and Validation

AlexNet-carpark - Mozilla Firefox

AlexNet-carpark DIGITS/GettingStart... localhost/models/20160718-081011-9f13

Search Version 3.0.0

DIGITS Image Classification Model

Learning Rate vs Epoch

Trained Models

Select Model: Epoch #30 Download Model

Test a single image

Image URL:

Upload image: cp1-243.jpg

Show visualizations and statistics

Classify One

Test a list of images

Upload Image List: No file selected. Accepts a list of filenames or urls (you can use your val.txt file)

Classify Many

Number of images use from the file: 100 Leave blank to use all

Number of images to show per category: 9

Top N Predictions per Category

System tray icons: Home, File, Network, Power, Software Updater, Date/Time: 08:16

Digits – Results

DIGITS - Mozilla Firefox

AlexNet-carpark DIGITS DIGITS/GettingStart... +

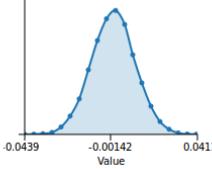
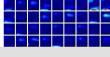
localhost/models/images/classification/classify_one?job_id=20160718-081011-9f13

Classify One Version 3.0.0

AlexNet-carpark Image Classification Model

Predictions

0	99.83%
1	0.17%

Description	Statistics	Visualization
data Activation	Data shape: (3, 51, 51) Mean: 8.40363 Std deviation: 26.9045	
conv1 Weights (Convolution layer) 34,944 learned parameters	Data shape: (96, 3, 11, 11) Mean: -0.000152168 Std deviation: 0.0102913	 
conv1	Data shape: (96, 11, 11) Mean: 2.13561 Std deviation: 0.0102913	

DIGITS - Mozilla Firefox [Software Updater] 08:17

Digits – Results

DIGITS - Mozilla Firefox

AlexNet-carpark DIGITS DIGITS/GettingStart... +

localhost/models/images/classification/classify_one?job_id=20160718-081011-9f13

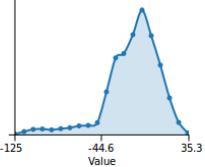
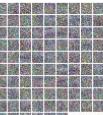
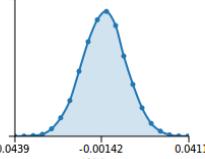
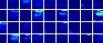
Search Version 3.0.0

DIGITS AlexNet-carpark Classify One

AlexNet-carpark Image Classification Model

Predictions

0	99.06%
1	0.92%

Description	Statistics	Visualization
data Activation	Data shape: (3, 51, 51) Mean: -17.4052 Std deviation: 24.4735	 
conv1 Weights (Convolution layer) 34,944 learned parameters	Data shape: (96, 3, 11, 11) Mean: -0.000152168 Std deviation: 0.0102913	 
conv1	Data shape: (96, 11, 11) Mean: 3.27513 Std deviation: 0.0102913	 

DIGITS - Mozilla Firefox [Software Updater] 08:18

Digits – Results

DIGITS - Mozilla Firefox

AlexNet-carpark DIGITS DIGITS DIGITS/GettingStart... +

localhost/models/images/classification/classify_one?job_id=20160718-081011-9f13

Search Version 3.0.0

DIGITS AlexNet-carpark Classify One

AlexNet-carpark Image Classification Model

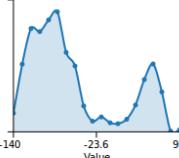
Predictions

1	99.99%
0	0.01%

Description Statistics Visualization

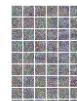
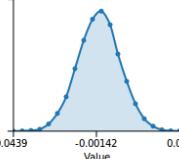
data
Activation

Data shape: (3, 51, 51)
Mean: -55.3716
Std deviation: 59.5018



conv1
Weights (Convolution layer)
34,944 learned parameters

Data shape: (96, 3, 11, 11)
Mean: -0.000152168
Std deviation: 0.0102913

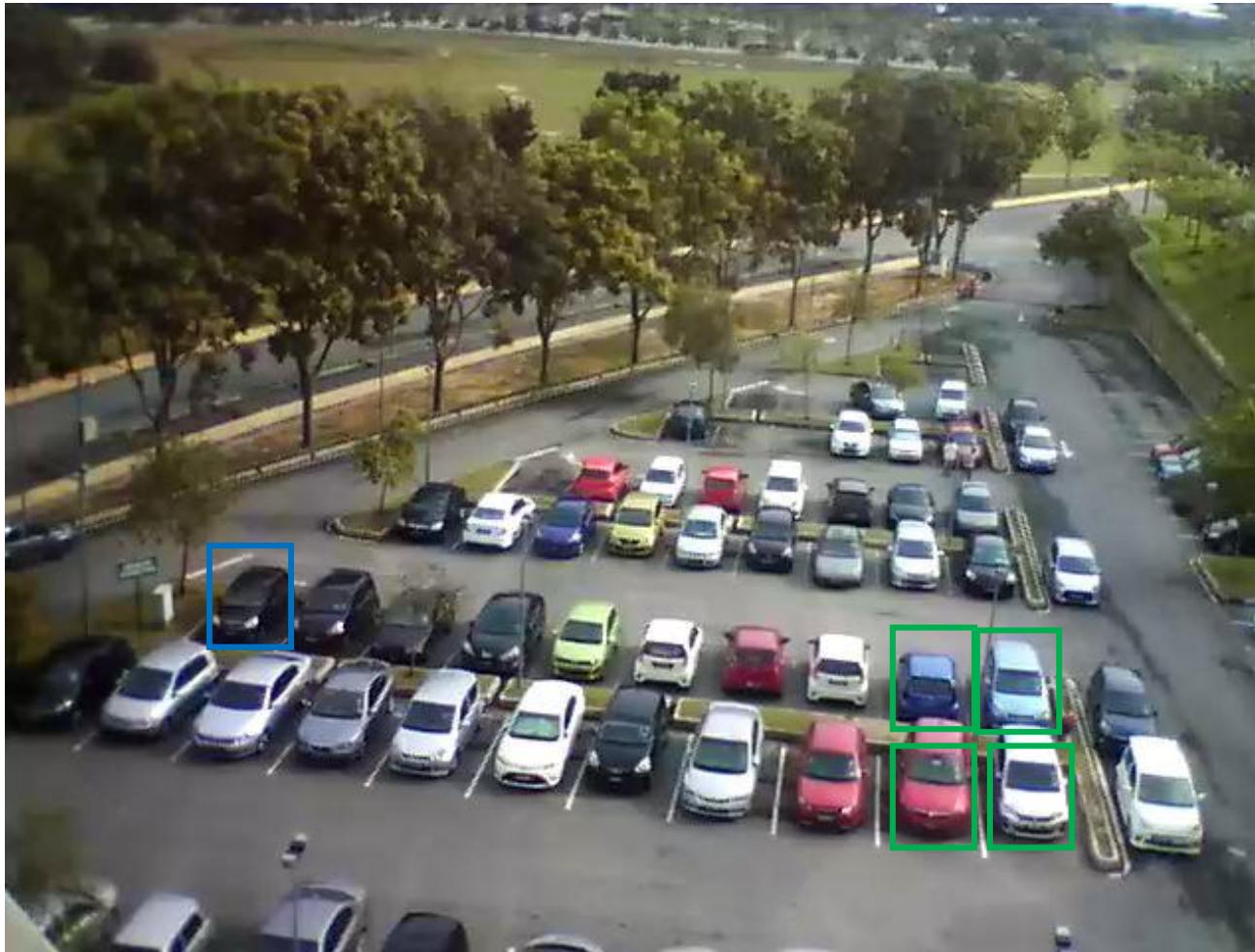


conv1

Data shape: (96, 11, 11)
Mean: 11.3505
Std deviation: 20.6666



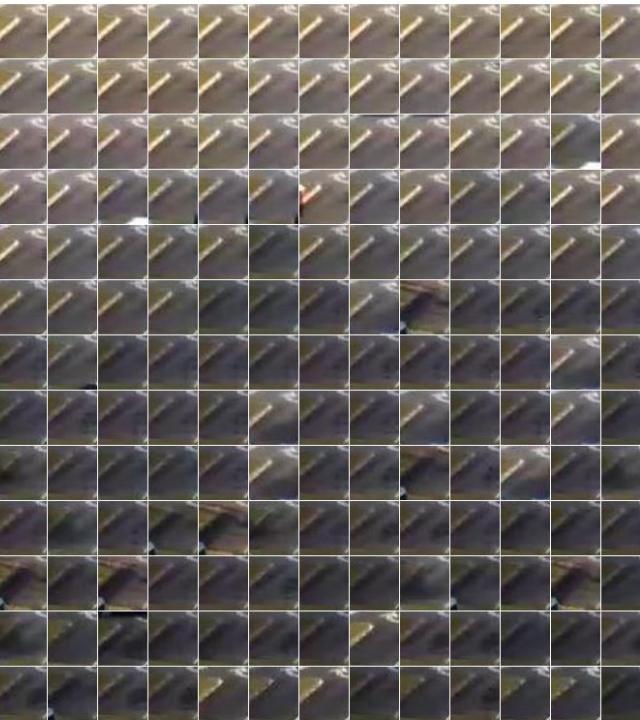
Digits – Further Evaluation



Digits – Further Evaluation

DIGITS Image Classification Model Top N

Top N Predictions per Category

Category	Top images for this category
0	

Fir... Software Updater screenshots

DIGITS Image Classification Model Top N

1











Digits – Further Evaluation

DIGITS - Mozilla Firefox

AlexNet-carpark DIGITS localhost/models/images/classification/ classify_many?job_id=20160718-081011-9f13

DIGITS AlexNet-carpark Classify Many Version 3.0.0

AlexNet-carpark Image Classification Model

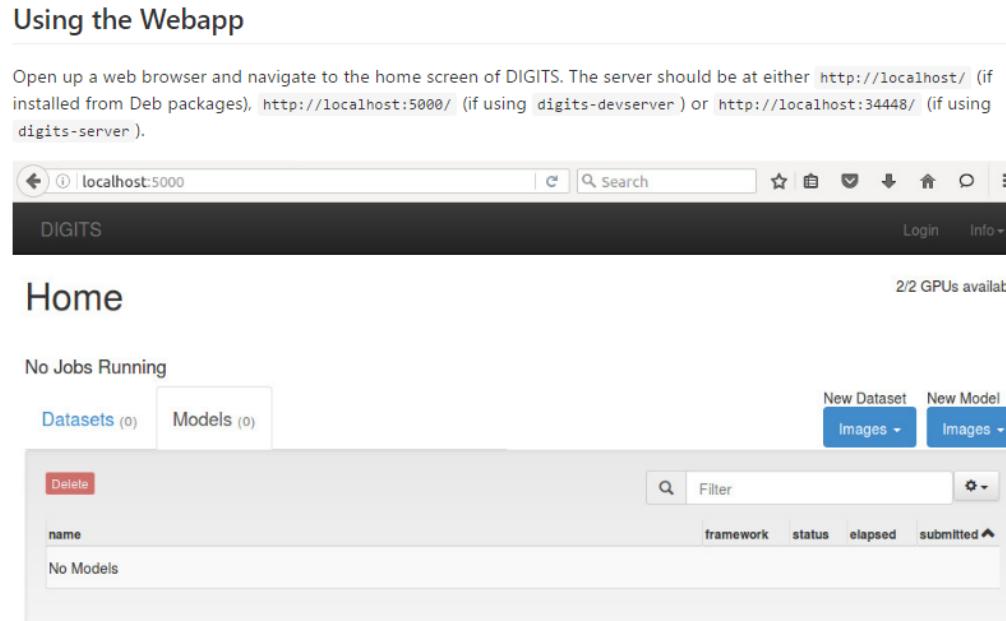
Path	Top predictions	
/home/hpc012/carpark/val/cp5-001.jpg	1 99.9%	0 (0.1%)
/home/hpc012/carpark/val/cp5-002.jpg	1 99.9%	0 (0.1%)
/home/hpc012/carpark/val/cp5-003.jpg	1 99.93%	0 (0.07%)
/home/hpc012/carpark/val/cp5-004.jpg	1 99.91%	0 (0.09%)
/home/hpc012/carpark/val/cp5-005.jpg	1 99.9%	0 (0.1%)
/home/hpc012/carpark/val/cp5-006.jpg	1 99.84%	0 (0.16%)
/home/hpc012/carpark/val/cp5-007.jpg	1 99.77%	0 (0.23%)
/home/hpc012/carpark/val/cp5-008.jpg	1 99.88%	0 (0.12%)
/home/hpc012/carpark/val/cp5-009.jpg	1 99.83%	0 (0.18%)
/home/hpc012/carpark/val/cp5-010.jpg	1 99.83%	0 (0.24%)
/home/hpc012/carpark/val/cp5-011.jpg	1 99.83%	0 (0.17%)
/home/hpc012/carpark/val/cp5-012.jpg	1 99.9%	0 (0.1%)
/home/hpc012/carpark/val/cp5-013.jpg	1 99.81%	0 (0.19%)
/home/hpc012/carpark/val/cp5-014.jpg	1 99.84%	0 (0.16%)
/home/hpc012/carpark/val/cp5-015.jpg	1 99.76%	0 (0.24%)
/home/hpc012/carpark/val/cp5-016.jpg	1 99.85%	0 (0.15%)
/home/hpc012/carpark/val/cp5-017.jpg	1 100.0%	0 (0.0%)
/home/hpc012/carpark/val/cp5-018.jpg	1 100.0%	0 (0.0%)
/home/hpc012/carpark/val/cp5-019.jpg	1 100.0%	0 (0.0%)
/home/hpc012/carpark/val/cp5-020.jpg	1 100.0%	0 (0.0%)

86% accuracy

DIGITS - Mozilla Firefox Software Updater screenshots 08:22

Digits – try it out yourself

- <https://github.com/NVIDIA/DIGITS/blob/master/docs/GettingStarted.md>



Logging in

Click on `New Dataset > Images > Classification`. This will lead you to the login page:

NOTE: there is no authentication - you don't even need a password. This is a utility feature, not a security feature.

Thanks!

Questions?

A large, colorful word cloud centered around the word "thank you". The word "thank you" is repeated in many different languages, each with its English translation. The languages include German (danke), Chinese (謝謝), Turkish (teşekkür ederim), Spanish (gracias), French (merci), Italian (grazie), Portuguese (obrigado), Polish (dziękuje), Russian (спасибо), Korean (감사합니다), Japanese (ありがとうございます), and many others. The words are in various colors and sizes, creating a dense and diverse visual representation of gratitude across cultures.



@kuanhhoong



<https://www.linkedin.com/in/kuanhoong>



khpoo@mmu.edu.my