OPEN DATA SCIENCE CONFERENCE

#ODSC

@ODSC

Boston | May 1 - 4 2018

TensorFlow Keras

POWER UP YOUR COMPUTER VISION SKILLS WITH TENSORFLOW-KERAS

Mo Patel @mopatel

#ODSC

ODSC East - Boston May 2018

http://bit.ly/odsckerascv

Introductions & Networking

Agenda

- Computer Vision Overview
- Overview of Keras & TensorFlow
- Getting started with Keras
- Anatomy of Keras Model
- Computer Vision Models using Keras
    - Image Classification
    - Transfer Learning
    - Object Detection
- Workign with Computer Vision Datasets
- Deploying Keras Model
- Computer Vision next…

Computer Vision Overview

Computer Vision 101

- Image Properties
  - Pixels
  - Dimension
  - Channels
  - Color Depth

Color Depth	Channel Range
1-Bit	0-1
8-Bit	0-255
16-Bit	0-65,535
24-Bit	0-16,777,215

Width (800 pixels)

Height (600 pixels)

8-bit Image

One Pixel

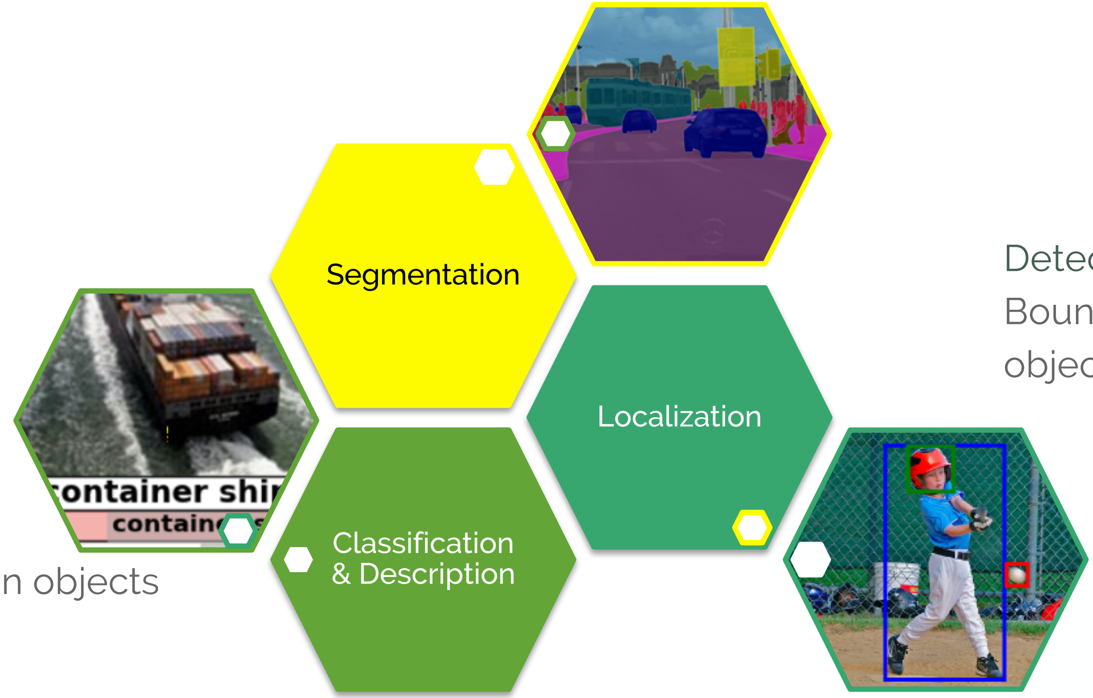
Channel	Red	Green	Blue
Value	0-255	0-255	0-255

Computer Vision 101

Segmentation: Pixel-wise segments of objects in any image

Classification: Label Item(s) in Image

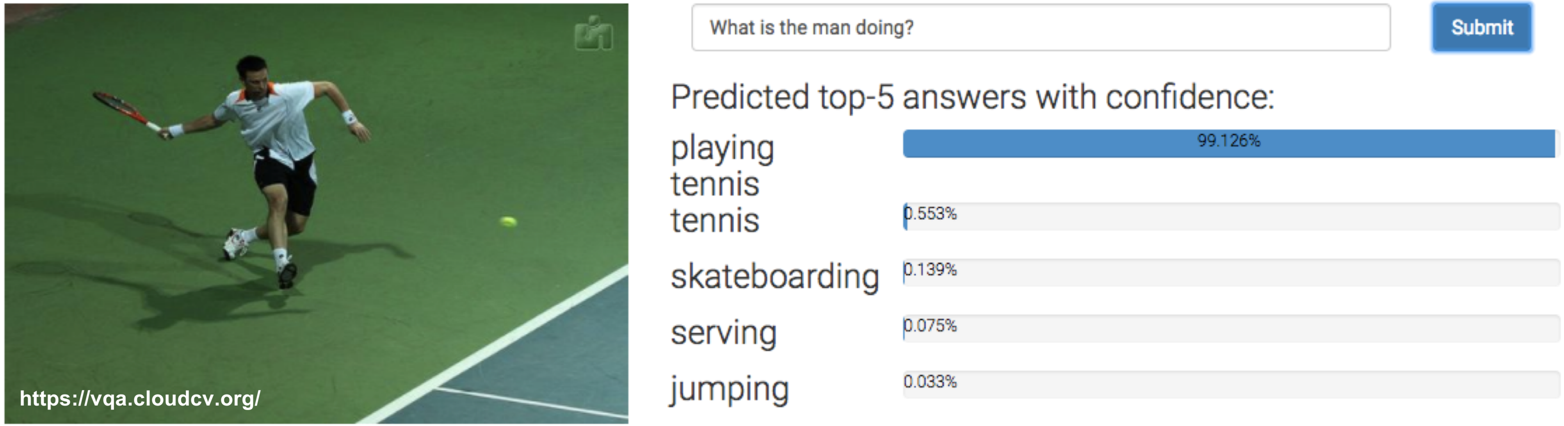
Description: Relationship between objects



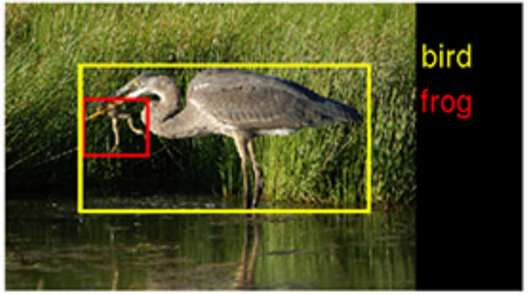
Detection/Localization: Bounding Boxes around objects

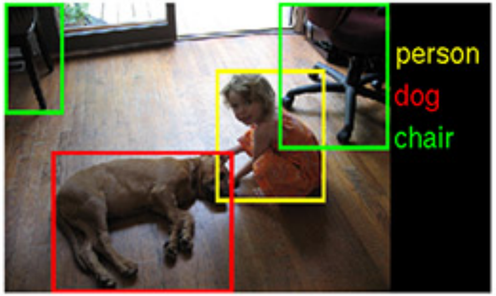
Goal: Build models that can perform visual tasks

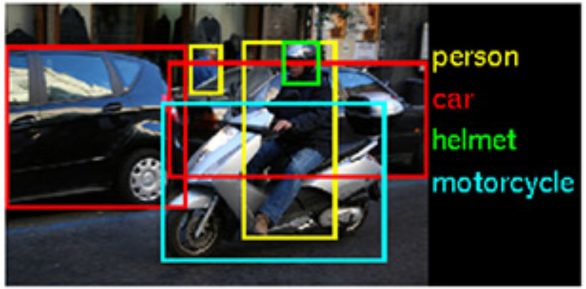
Classification & Description



Localization





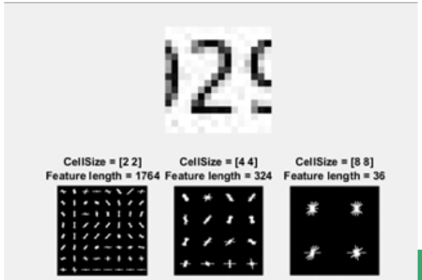


Source: Kaggle ImageNet Object Detection Challenge

Computer Vision prior to Deep Learning

- Objectives
  - Extract generalized features (patterns) from images
  - Determine generalized descriptors for feature
  - Use generalized features to perform visual tasks based on descriptors

- What are features?
  - Edges
  - Shapes
  - Change in Colors/Intensity



Histogram of oriented Gradients

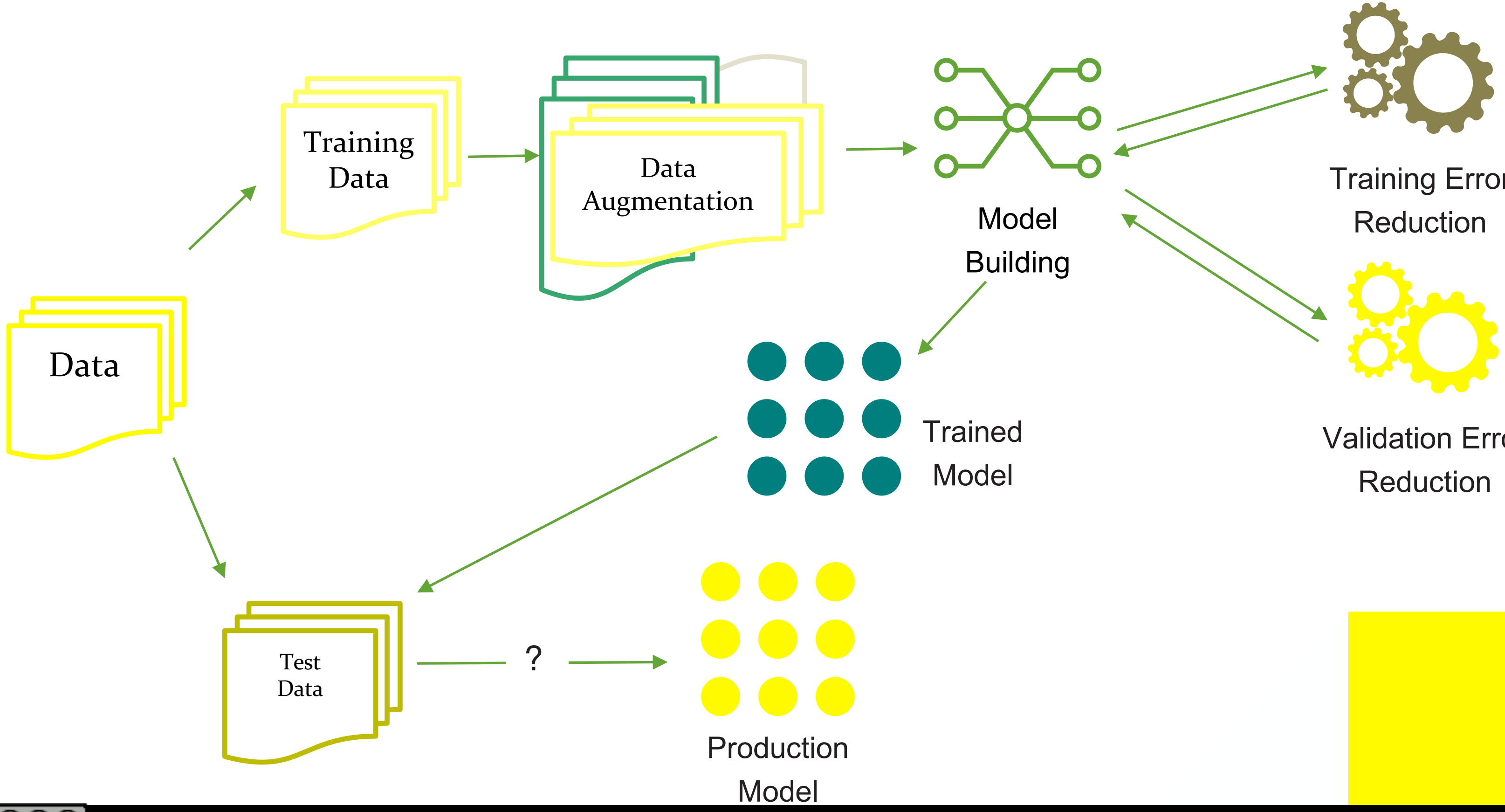
Traditional Machine Learning → Deep Learning

- Featuring engineering challenges
  - Human scaling: Time consuming to create rules for each object
  - Computational scaling: Translating human curated features to algorithms that machines can perform at scale is tedious
- Feature Learning
  - Design computational models to learn features
- Feature Learning Challenges
  - Need large datasets with annotations
  - Model architectures become complex both to design and run

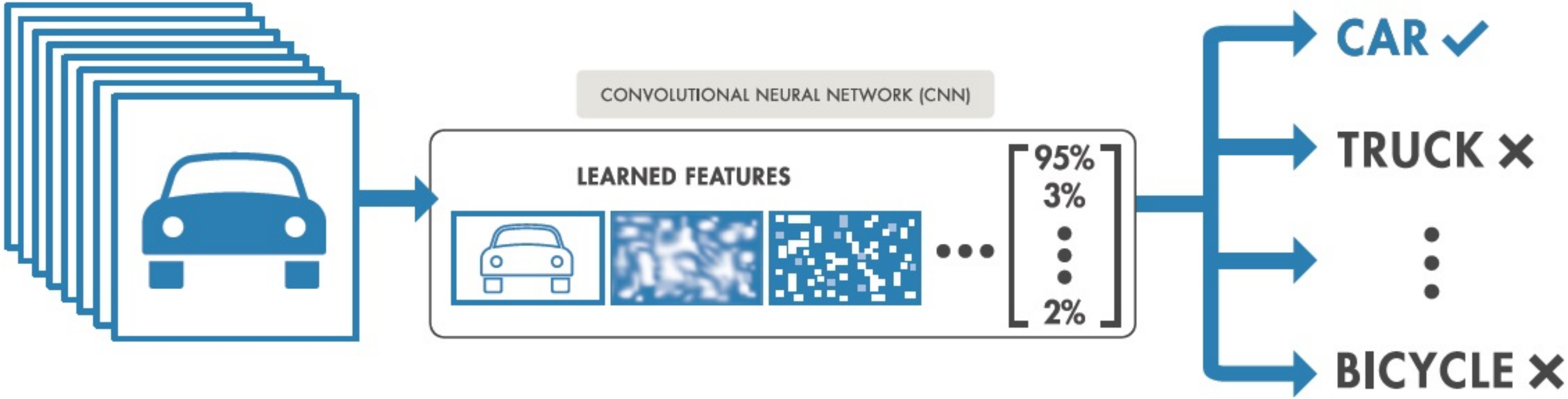
Convolutional Neural Networks

http://bit.ly/odsckerascv

Deep (Supervised Machine) Learning Basics



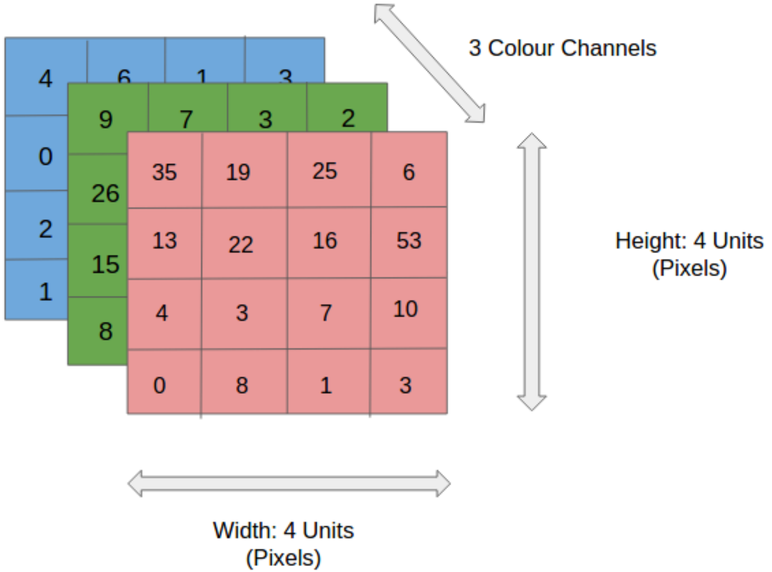
Convolutional Neural Networks



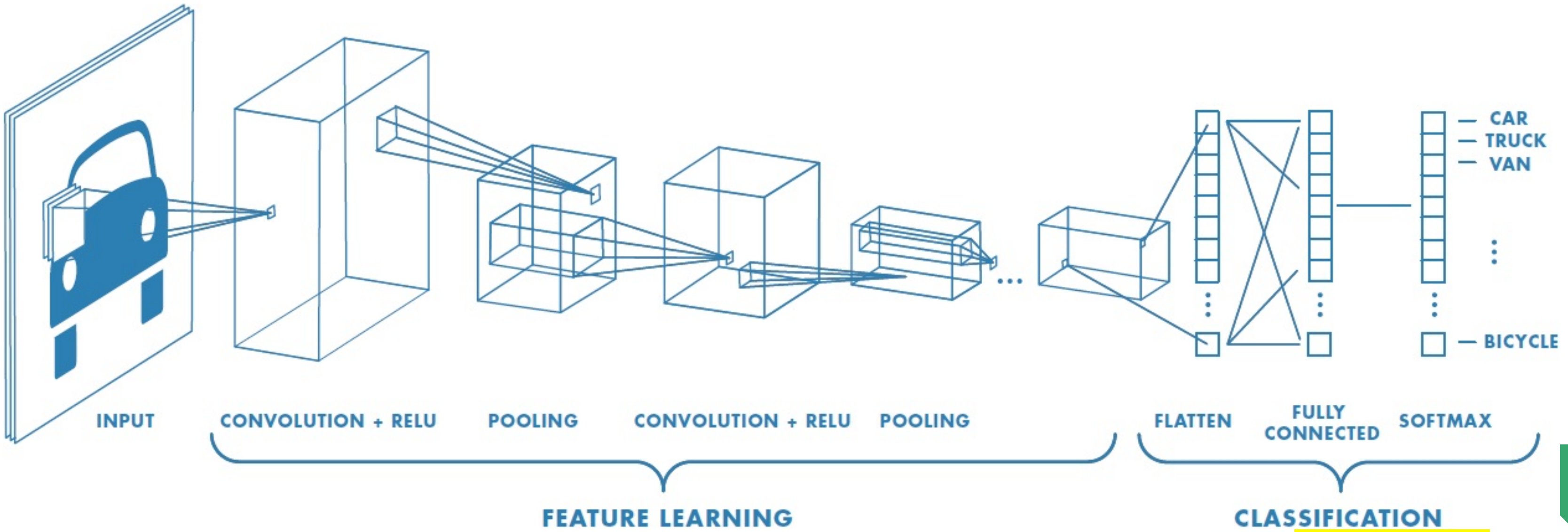
Source: Mathworks

Convolutional Neural Network Properties

- Learning Low Level Features from High Dimensional Data
- Must Have Data Properties:
  - Shift Invariant
  - Correlated
- Simply put: Changing column order should not change meaning of Data
- Naturally Suitable Data Types:
  - Images, Audio, Video
  - Time Series

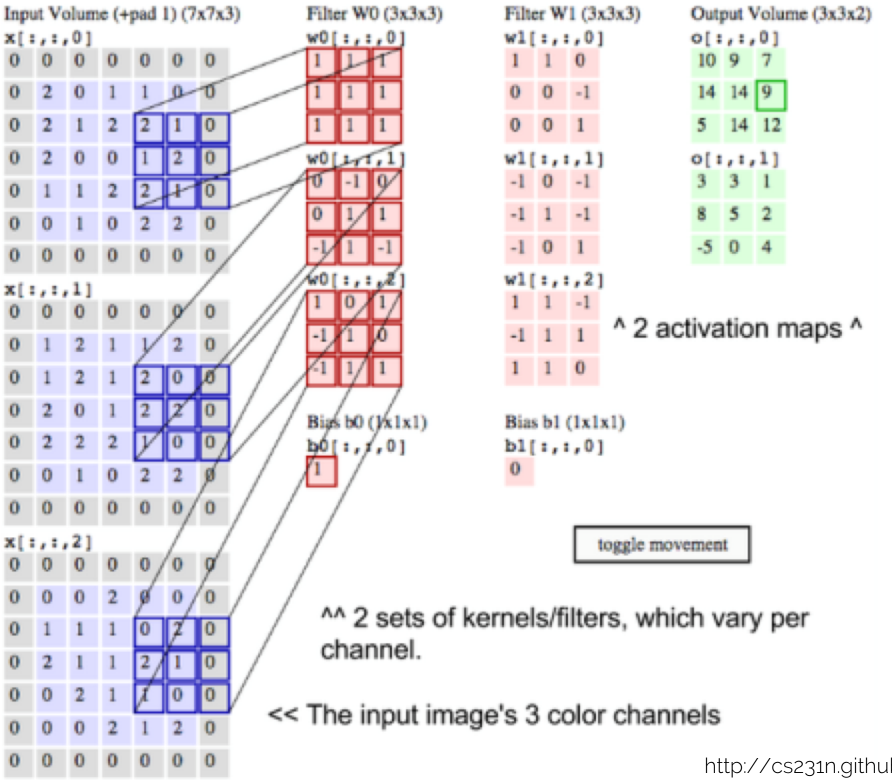


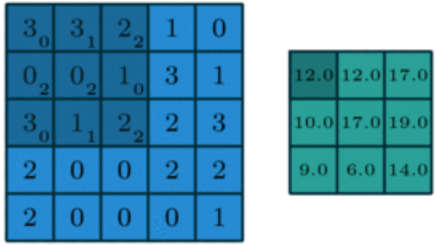
Convolutional Neural Networks



Source: Mathworks

Convolution

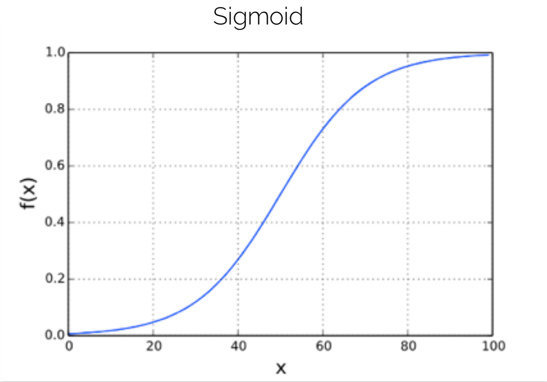


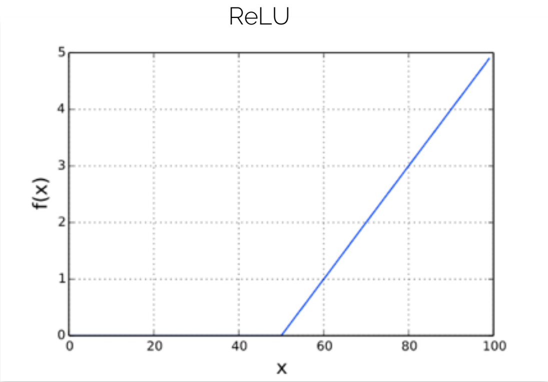


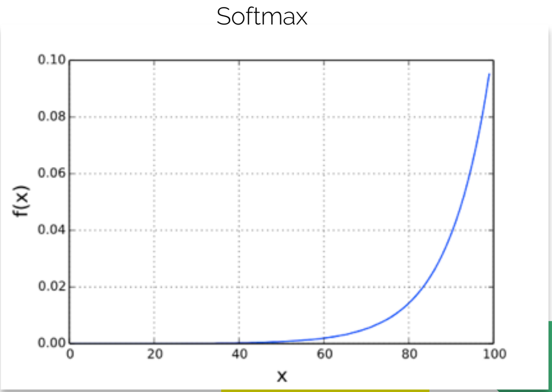
http://cs231n.github.io/convolutional-networks/
https://m2dsupsdlclass.github.io/lectures-labs/slides/03_conv_nets/index.html
https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

Activation

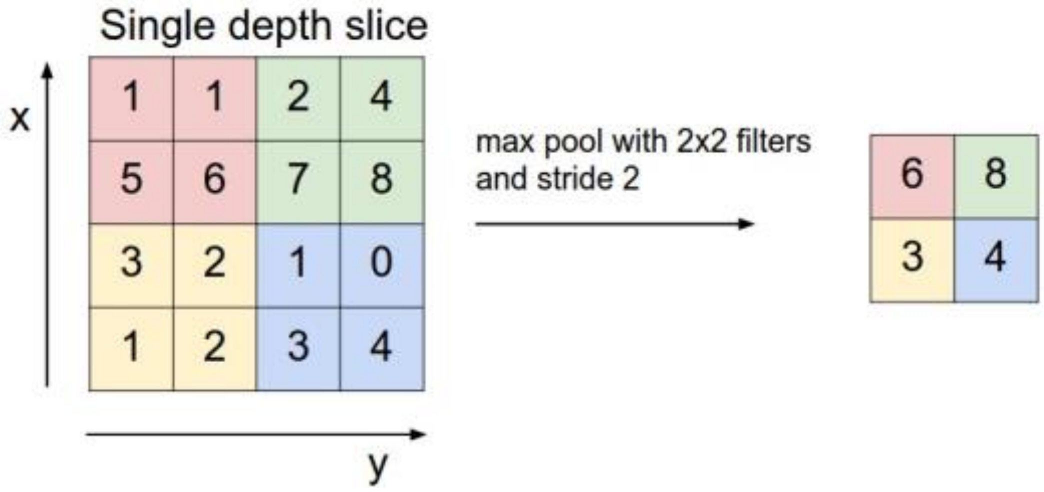
- Mechanism that decides to activate or not activate a neuron in a neural network







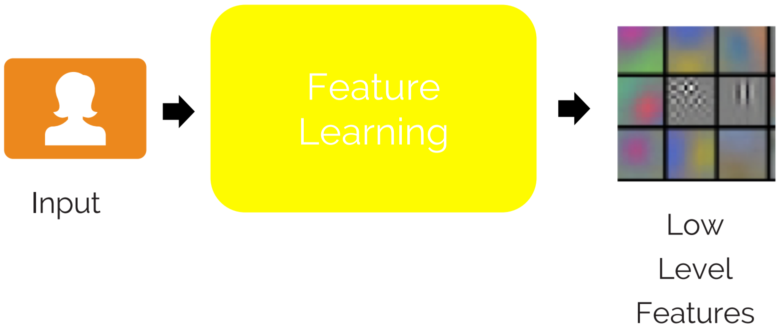
Pooling



Source: Stanford University CS 231 Course:
http://cs231n.github.io/convolutional-networks/#pool

Fully Connected

- Traditional Machine Learning: Feature Engineering
- Deep Learning: Feature Learning
- Learned Features are fed to traditional classifier such as Multivariate Logistic Regression (Softmax)



Flatten
10
24
2
3
...
...
...

Class	Probability
A	0.01
B	0.07
C	0.13
D	0.60
E	0.03
F	0.05
G	0.11

Loss Reduction

Class	Probability	Truth	Loss/ Error
A	0.01	0	0
B	0.07	0	0
C	0.13	0	0
D	0.60	1	0.22
E	0.03	0	0
F	0.05	0	0
G	0.11	0	0

Cross Entropy Loss

One-Hot Encoded Vector of Class Label

Back Propagation

- Account for loss across the weights and biases in the network
- Use cost function provided by Back Propagation to optimize the weights and biases to reduce loss
- Variety of Stochastic Gradient Descent techniques used for optimization

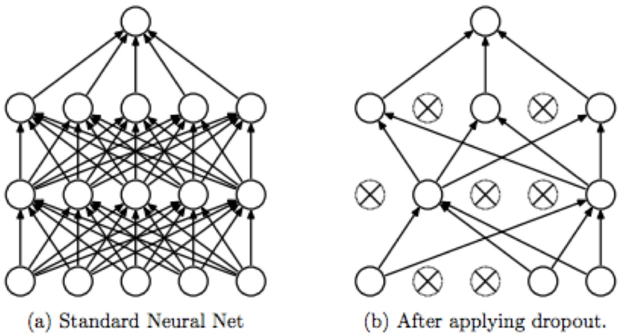
Overview of Gradient Descent Techniques:
http://ruder.io/optimizing-gradient-descent/

https://rdipietro.github.io/friendly-intro-to-cross-entropy-loss/

Generalization

Dropout

Randomly remove learned nodes from network



Dropout: A Simple Way to Prevent Neural Networks from Overfitting
http://www.jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf

Batch Normalization

- During training input data shifts as weights and parameter adjust values
- To prevent internal covariate shift, normalize each batch by mean and variance

Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift
https://arxiv.org/abs/1502.03167