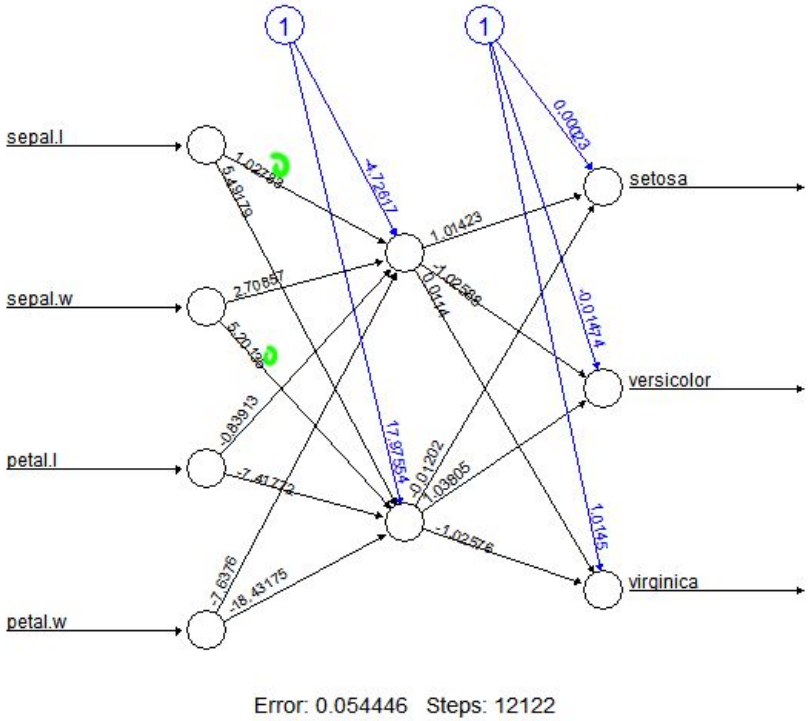
Week 8

Agenda

1. Project 4 check-in
2. NN recap
3. Convolutional Networks
4. Breakout

Example Trained Neural Network



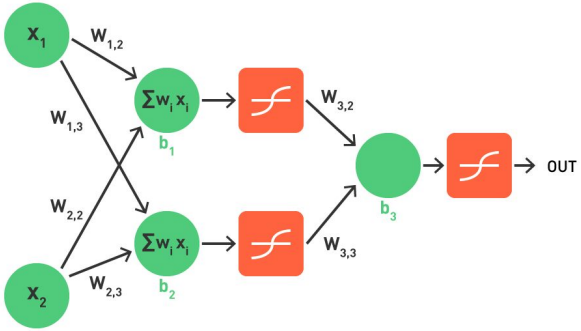
1. What are the parameters in this NN?
2. Is this a sparse or dense NN?
3. Advantages/disadvantages of sparse vs dense NN?
4. How can we think about this network as an ensemble/stacked model?
5. How can we think about this network as a series of matrix operations?

Neural Network Recap

1. What happens in forward propagation?
2. What happens in backpropagation?
3. What are the benefits of SGD and Mini-batches?

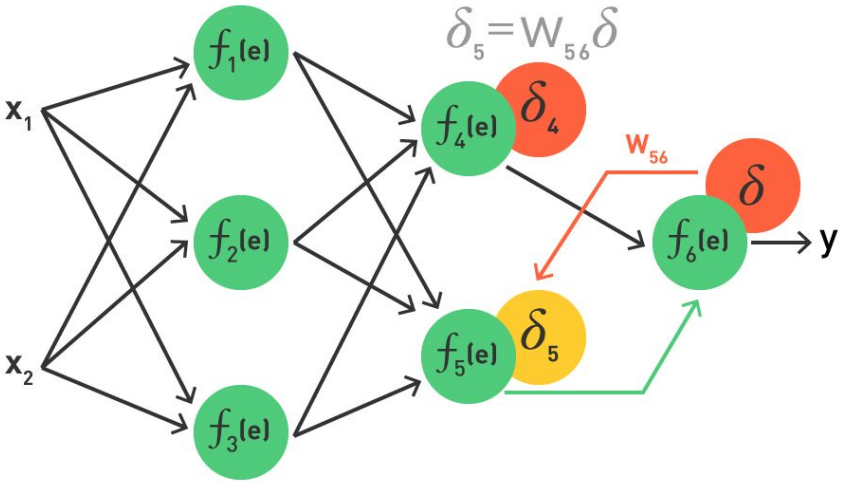
Intuition: Forward Propagation

- Given a training example (X_1, X_2) and output Y_i
- Propagate inputs/activations forward, applying sigmoid function on dot products

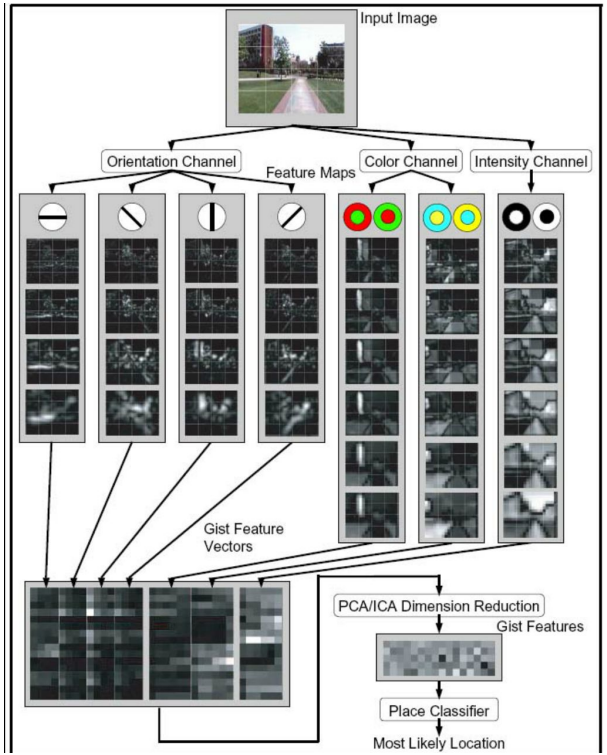


Intuition: Backward Propagation (cont.)

Propagate costs backward to earlier nodes:



Computer Vision (CV)



Conferences

- CV is discussed at most ML and AI conferences
- CVPR is main CV conference

Datasets

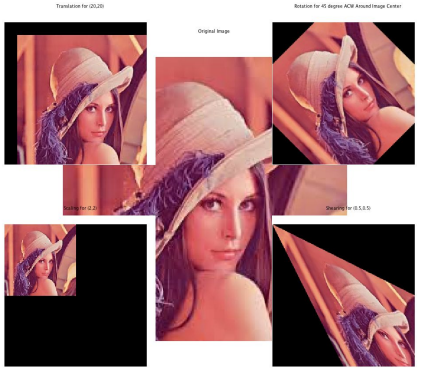
- MNIST -- 60,000 images
- SVHN (http://ufldl.stanford.edu/housenumbers/) -- 600,000 images
- ImageNet (http://image-net.org/about-stats) -- 14M images, 1TB, mapped to WordNet, includes features and hand labels

Feature Engineering

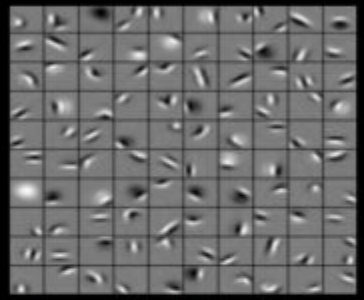
- A major focus of field
- SIFT: 1999, patented by BC.
- Also SURF, GIST, HOG

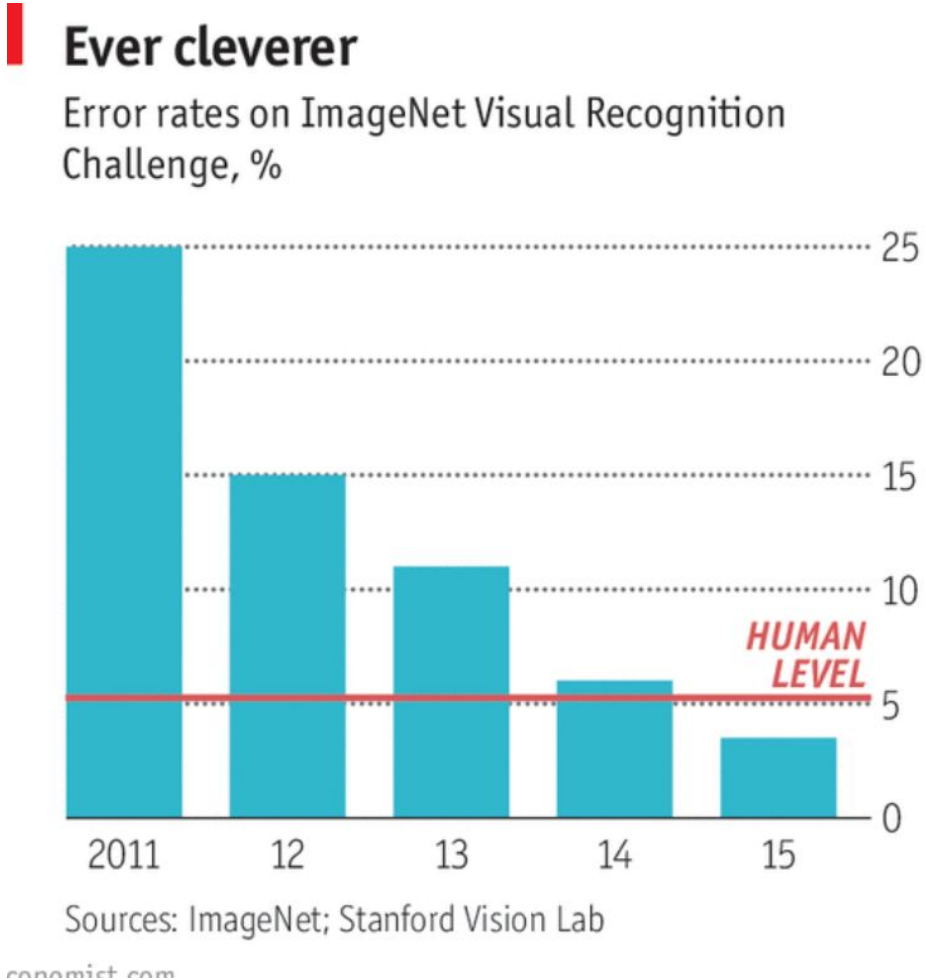
Engineering Examples

- Common in CV
- Do things that maintain label: Rotate, translate, skew, scale, etc



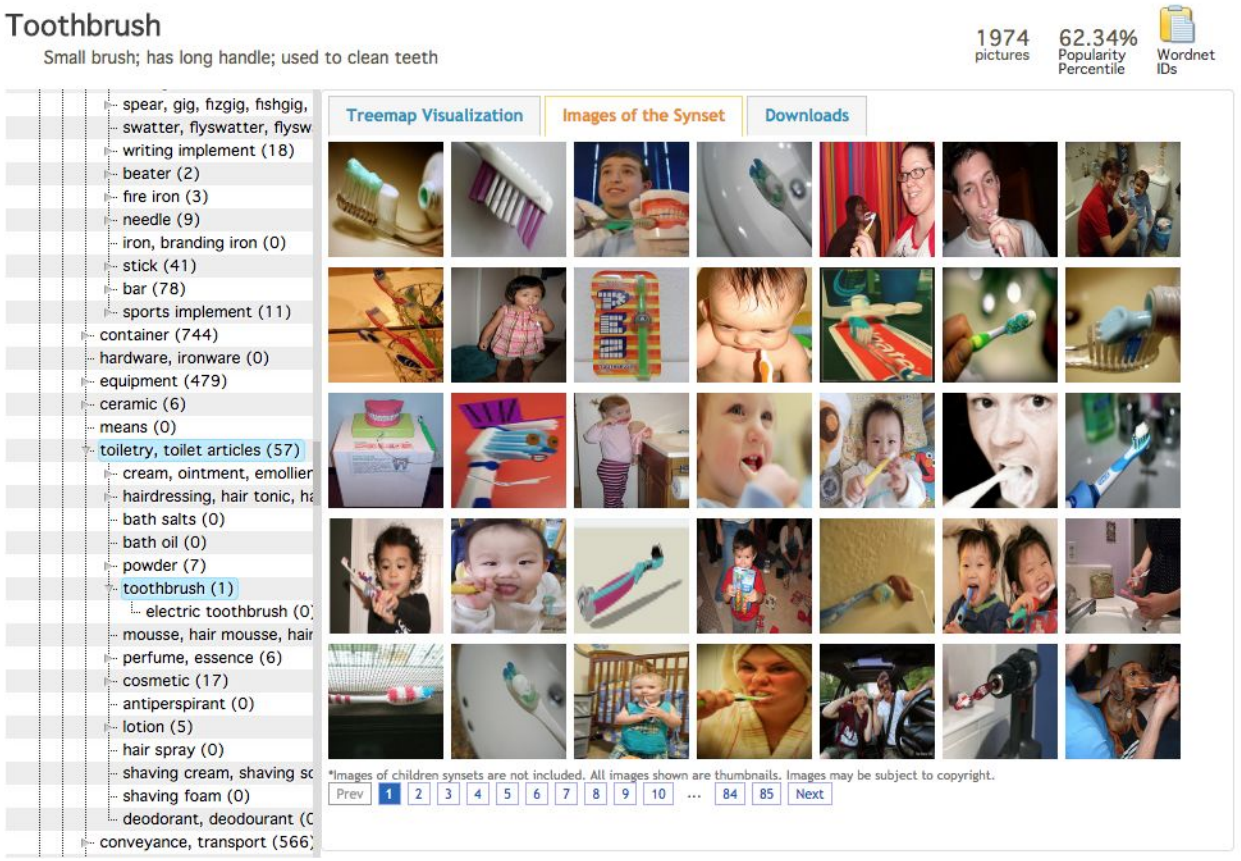
Deep Learning learns layers of features

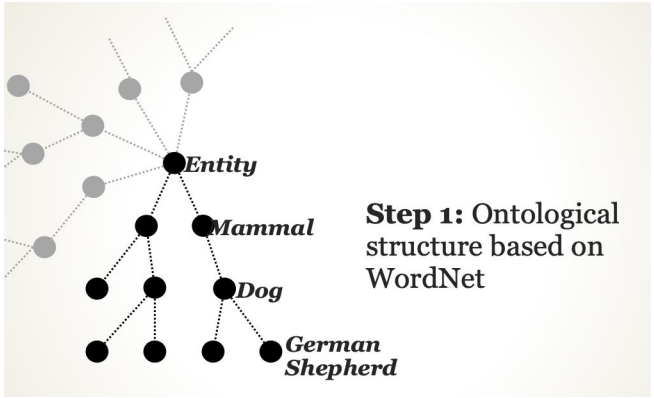




Deep (Feature/Representation) Learning

- Move away from feature engineering (still some and some Architectural design)
- Today learned features generally outperform
- Learn similar gradient based features at early layers


Toothbrush
Small brush; has long handle; used to clean teeth
1974 pictures
62.34% Popularity Percentile
Wordnet IDs
spear, gig, fizgig, fishgig,
swatter, flyswatter, flysw
writing implement (18)
beater (2)
fire iron (3)
needle (9)
iron, branding iron (0)
stick (41)
bar (78)
sports implement (11)
container (744)
hardware, ironware (0)
equipment (479)
ceramic (6)
means (0)
toiletry, toilet articles (57)
cream, ointment, emollier
hairdressing, hair tonic, ha
bath salts (0)
bath oil (0)
powder (7)
toothbrush (1)
electric toothbrush (0)
mousse, hair mousse, hair
perfume, essence (6)
cosmetic (17)
antiperspirant (0)
lotion (5)
hair spray (0)
shaving cream, shaving sc
shaving foam (0)
deodorant, deodourant (C
conveyance, transport (566)
Treemap Visualization
Images of the Synset
Downloads
*Images of children synsets are not included. All images shown are thumbnails. Images may be subject to copyright.
Prev 1 2 3 4 5 6 7 8 9 10 ... 84 85 Next



Step 1: Ontological structure based on WordNet

Step 2: Populate categories with thousands of images from the Internet

Step 3: Clean results by hand

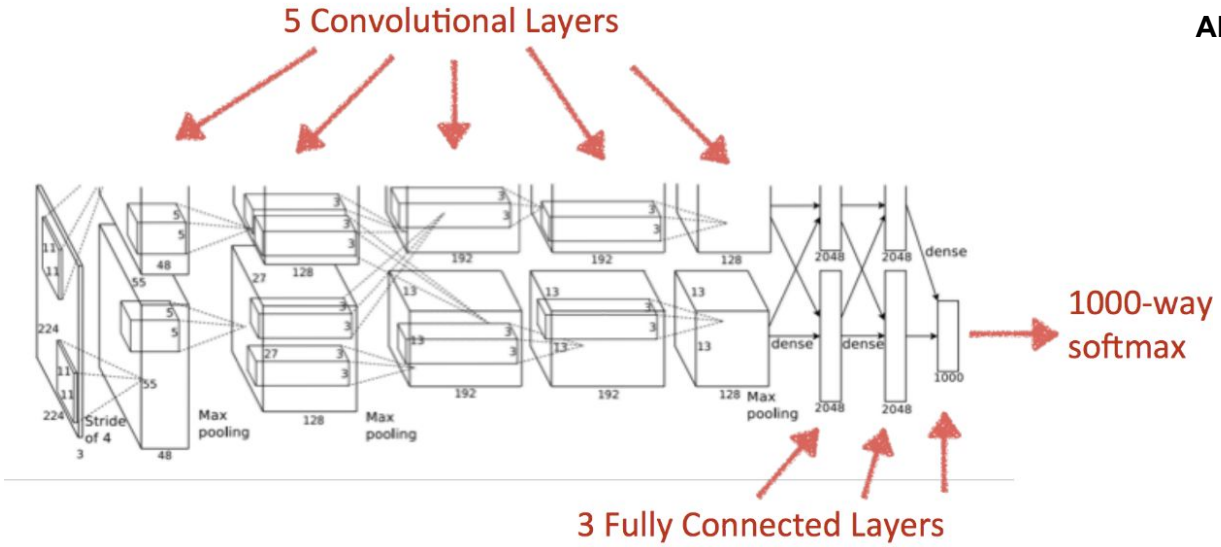
ImageNet
PhD Students

Crowdsourced Labor

amazon mechanical turk™
Artificial Artificial Intelligence

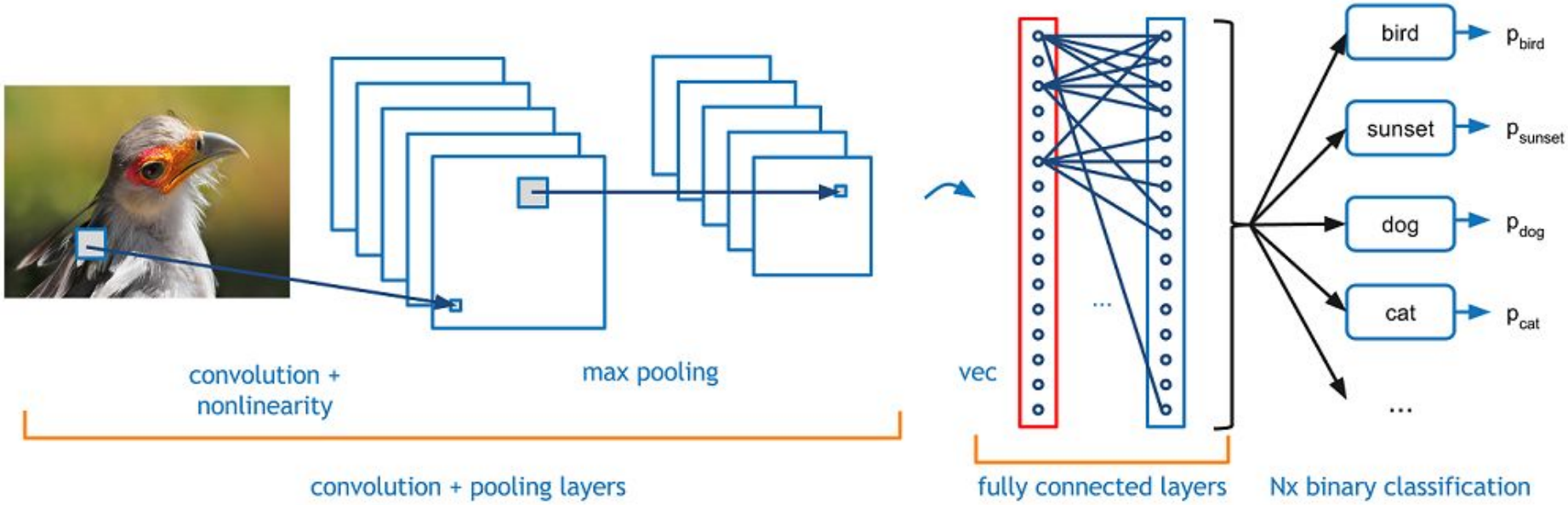
49k Workers *from* 167 Countries
2007-2010

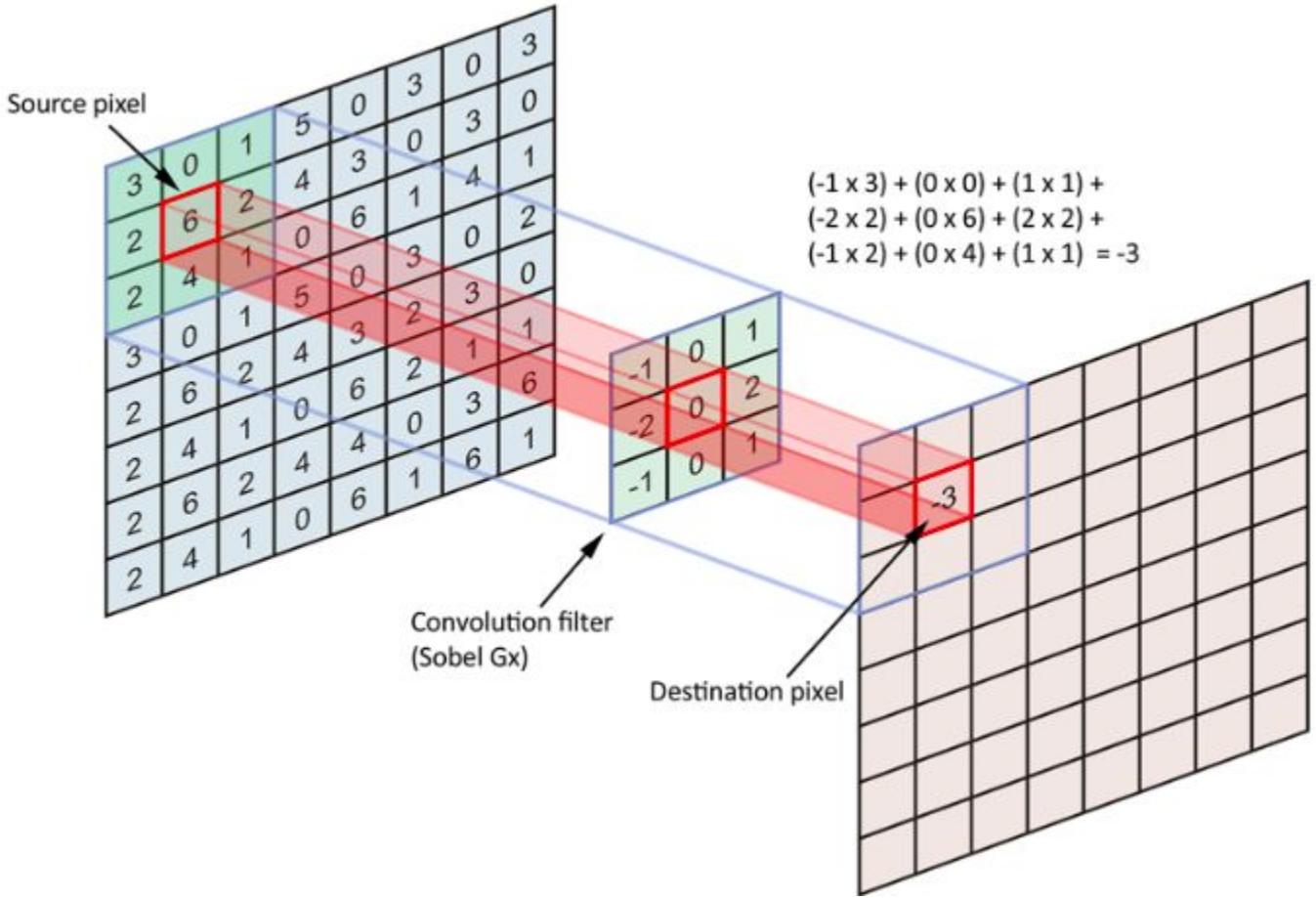
Convolution Nets

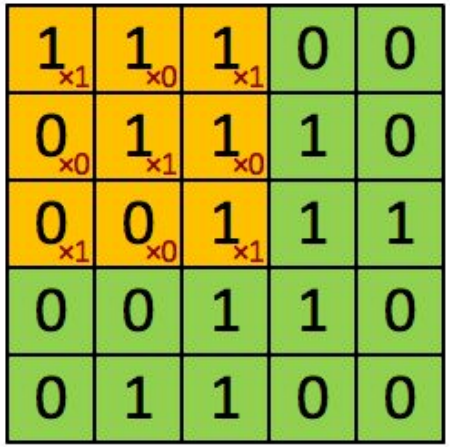


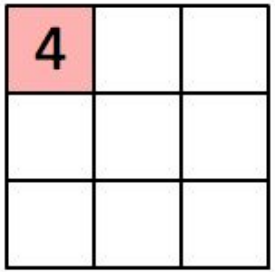
About

- Yann LeCun. LeNet: http://yann.lecun.com/exdb/lenet/ (1989-1998)
- Inspired by Visual Cortex in cats (receptive fields)
- Designed with image recognition in mind--input and layers often shown as 2D or 3D which may look odd coming from 1D.
- Composition of layers. Rightmost feature layers are most similar to output in representation
- Feature learning layers are of different types: (1) convolution and (2) pooling
- AlexNet 2012 (with Hinton) (picture)
- (https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf)


bird
sunset
dog
cat
...
p_{bird}
p_{sunset}
p_{dog}
p_{cat}
convolution + nonlinearity
max pooling
vec
convolution + pooling layers
fully connected layers
Nx binary classification


Source pixel
(-1 x 3) + (0 x 0) + (1 x 1) +
(-2 x 2) + (0 x 6) + (2 x 2) +
(-1 x 2) + (0 x 4) + (1 x 1) = -3
Convolution filter
(Sobel Gx)
Destination pixel
-3





gif source:
https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Visualization of the filter on the image

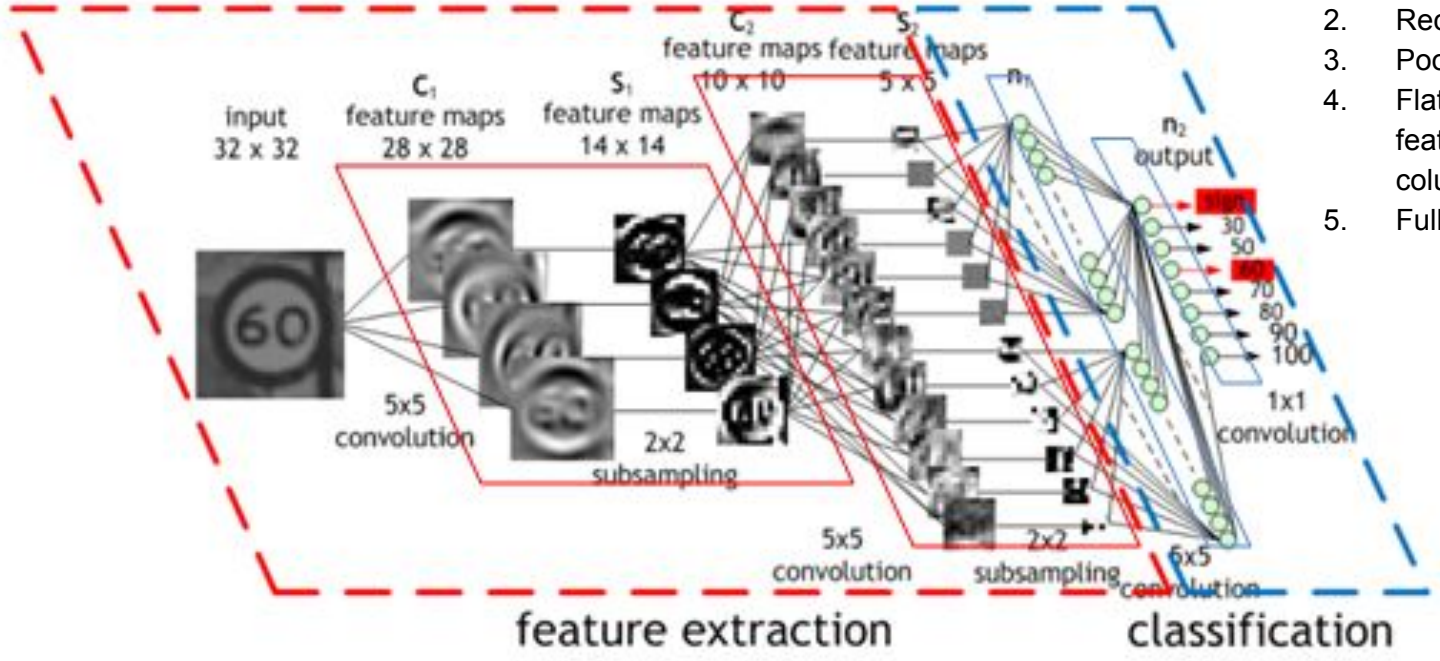
0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

Pixel representation of receptive field

*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

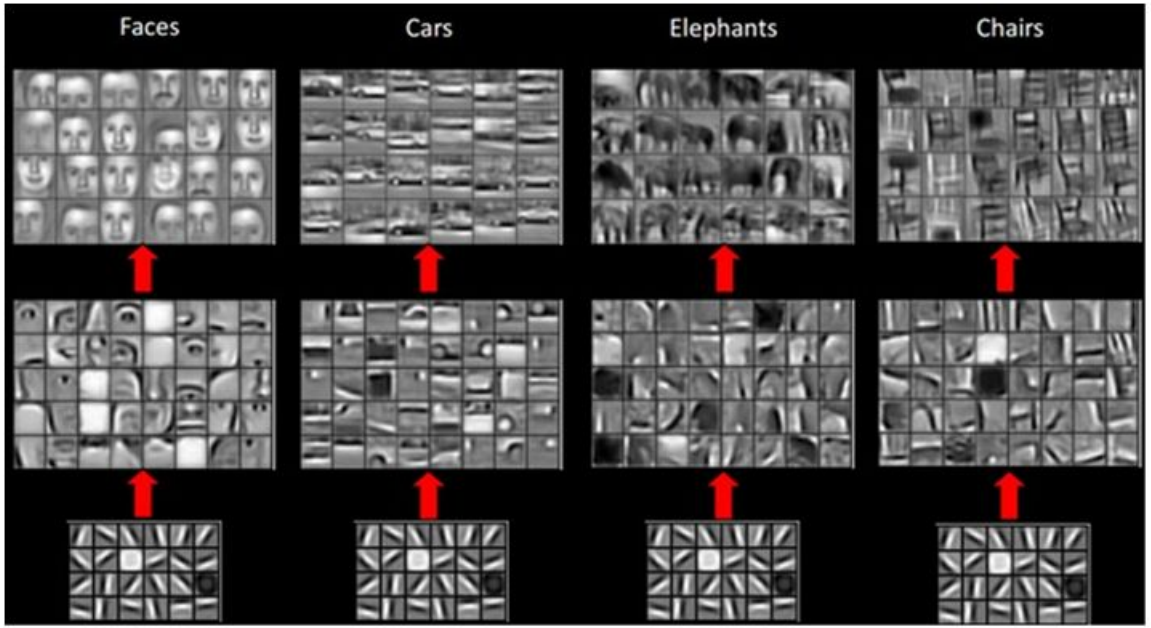


Training

1. Convolution
2. Rectifier function (ReLu)
3. Pooling
4. Flattening (transforming the pooled feature map matrix into a single column)
5. Fully connected

Breakout

1. What is a kernel? How does convolution work? http://setosa.io/ev/image-kernels/
2. Explain Pooling: what does it do? https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/
3. What is flattening? What does the fully connected part do?



Transfer Learning

- Train on one task, and use trained network or part of trained network when training for a different task
- Model Zoo (e.g. http://caffe.berkeleyvision.org/model_zoo.html)
- https://www.kaggle.com/c/state-farm-distracted-driver-detection/forums/t/20141/official-pre-trained-models-and-external-data-thread/116805