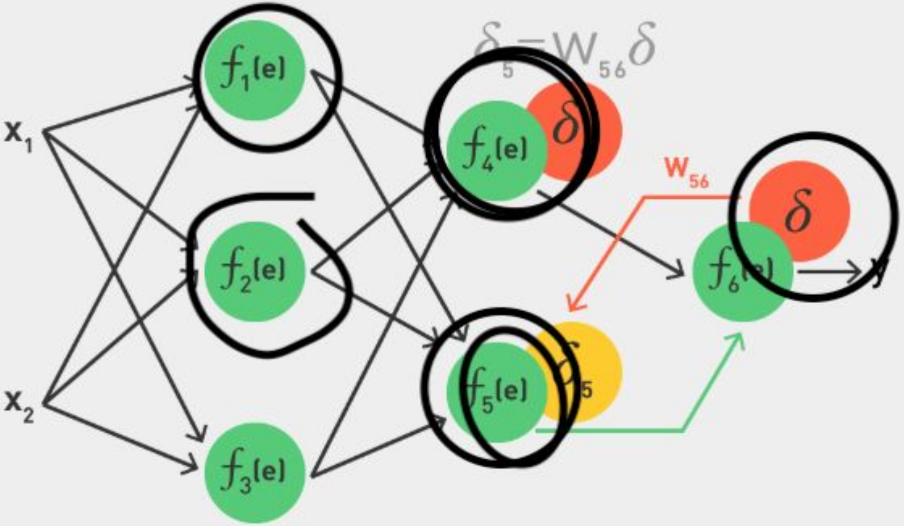
Week 8

Agenda

1. Review Neural Networks
2. Computer Vision discussion
3. LeCun paper discussion
4. Convolutional Networks introduction
5. Finish deep learning notebook
6. SVM discussion

Intuition: Backward Propagation (cont.)

Propagate costs backward to earlier nodes:



- For each hidden unit h in k^{th} layer:

$$\delta_{hk} = Y_{hk}\left(1 - Y_{hk}\right) \sum_{j \in K} w_{hj}\delta_j$$

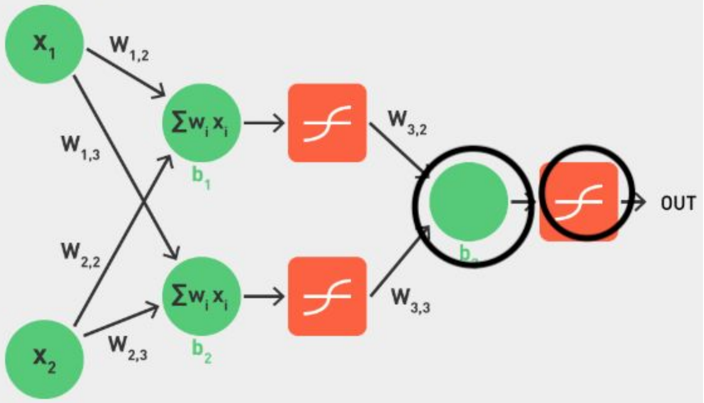
- Update each weight as $+\eta\delta_{hk}x_i$.
  - Daume ch. 8 for full algorithm

Neural Network Recap

1. What happens in forward propagation?
2. What happens in backprop?
3. What are the benefits of SGD and Mini-batches?
4. Why do GPUs speed up computation?
5. How do we handle regression problem?

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



Halle Berry Neuron

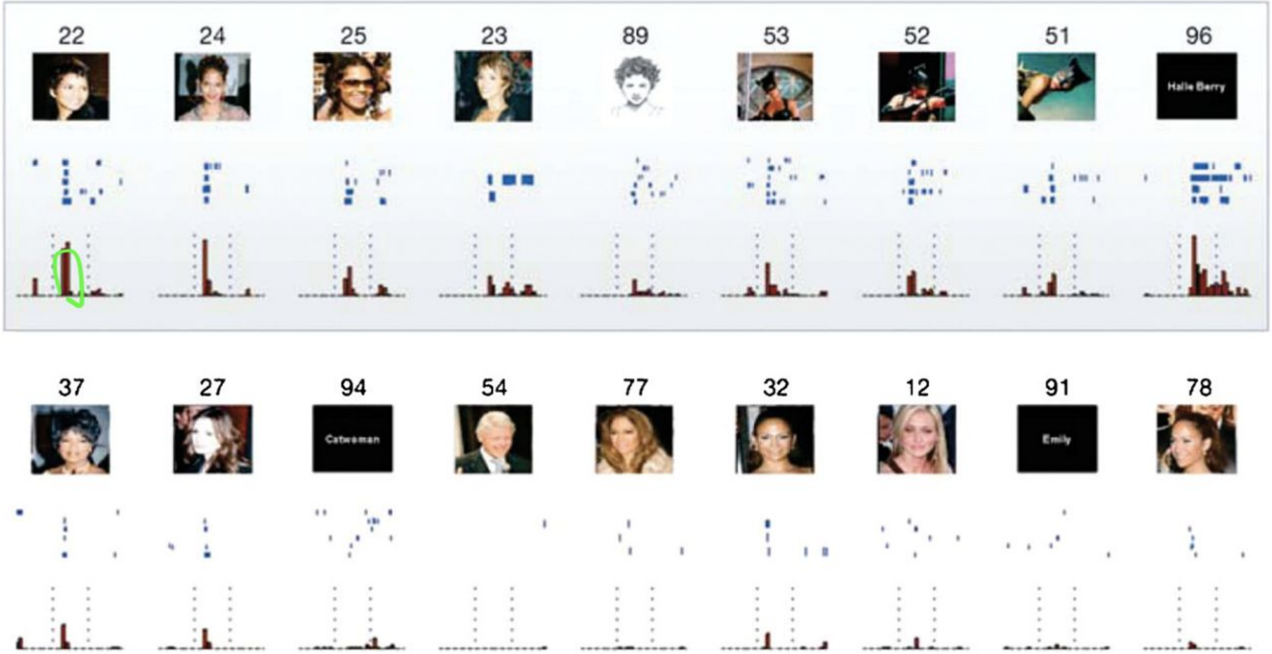
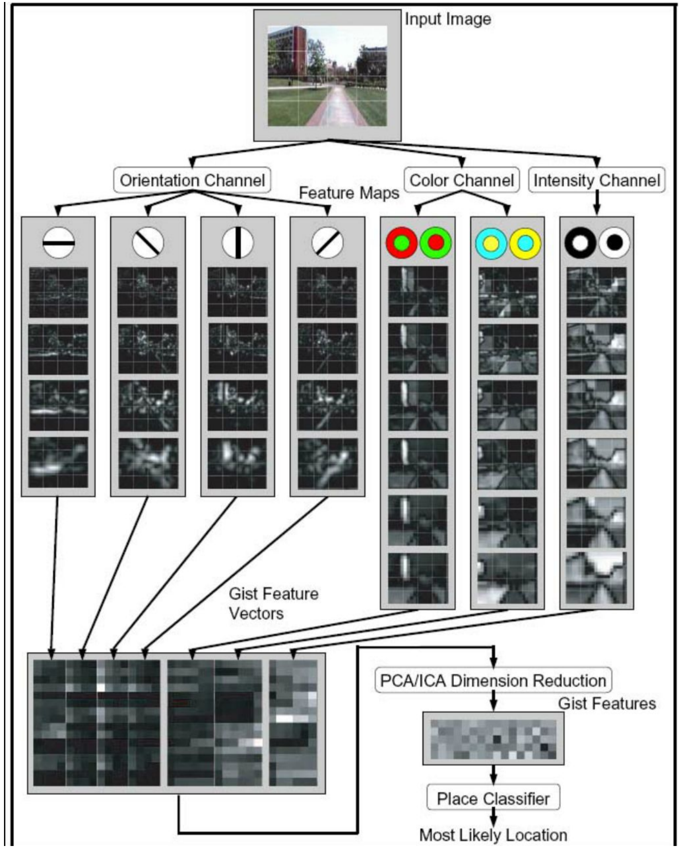


Fig. 8. A single unit in the human right anterior hippocampus that responds to different pictures of the actress Halle Berry including in costume and to the letter string of her name but not to other facial images or letter strings (Quiroga et al., 2005).

Computer Vision



Conferences

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

Datasets

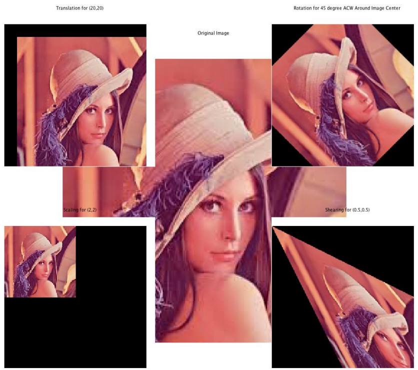
- MNIST -- 60,000 images
- SVHN (http://ufldl.stanford.edu/housenumbers/) -- 600,000 images
- ImageNet (http://image-net.org/about-stats) -- 14M images, 1TB, mapped to WordNet, includes features and hand labels

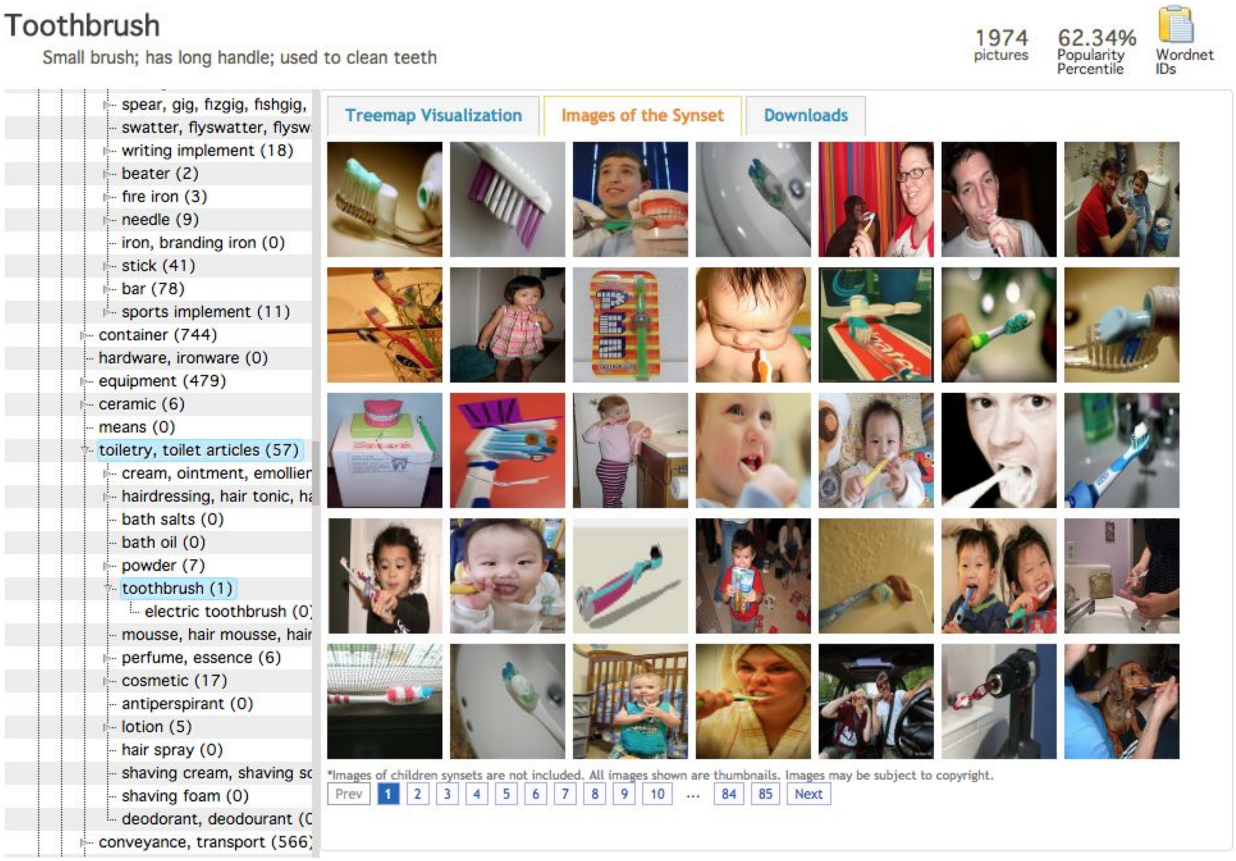
Feature Engineering

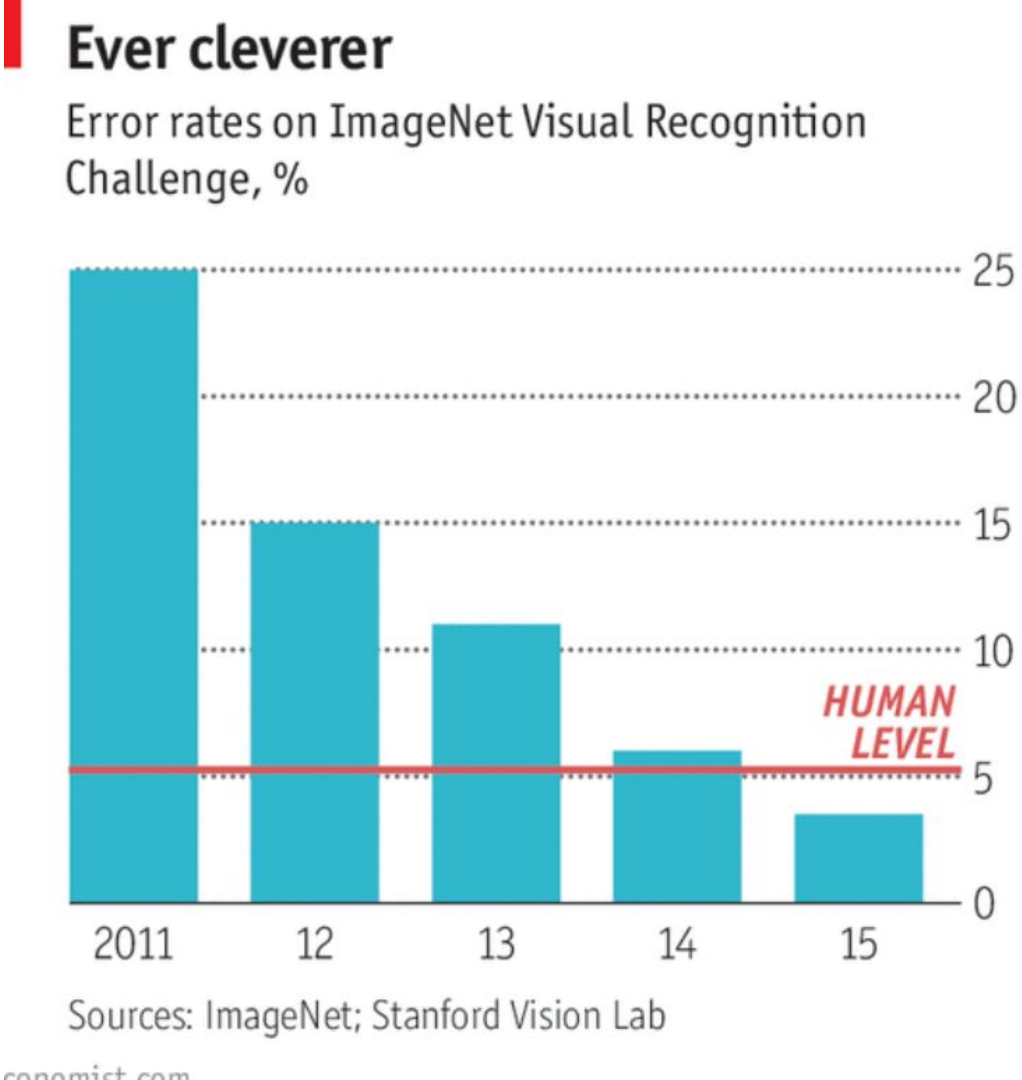
- A major focus of field
- Gradient based features popular
- SIFT: 1999, patented by BC.
- Also SURF, GIST, HOG

Engineering Examples

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



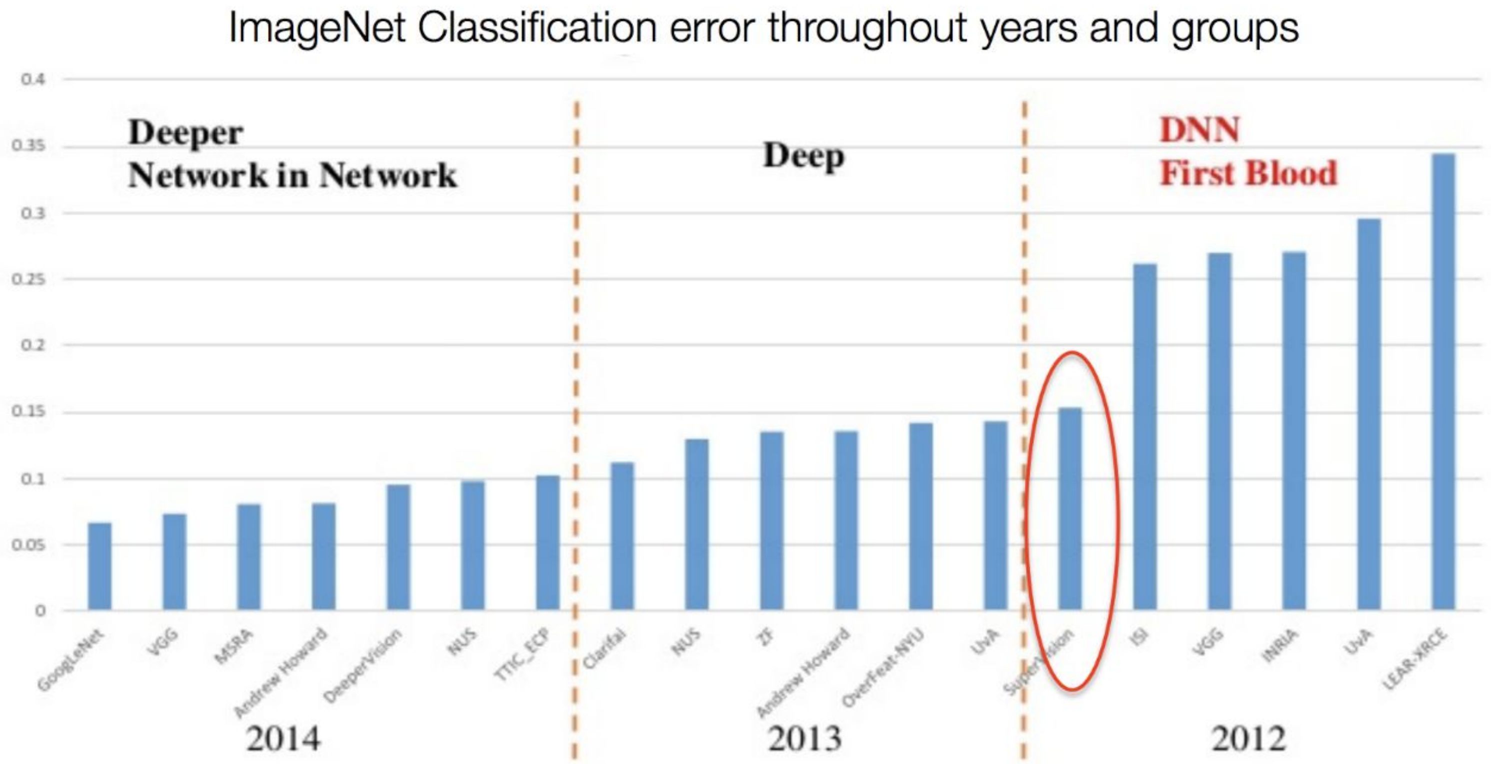

Toothbrush
Small brush; has long handle; used to clean teeth
1974
pictures
62.34%
Popularity Percentile
Wordnet IDs
spear, gig, fizgig, fishgig,
swatter, flyswatter, flysw
writing implement (18)
beater (2)
fire iron (3)
needle (9)
iron, branding iron (0)
stick (41)
bar (78)
sports implement (11)
container (744)
hardware, ironware (0)
equipment (479)
ceramic (6)
means (0)
toiletry, toilet articles (57)
cream, ointment, emollien
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



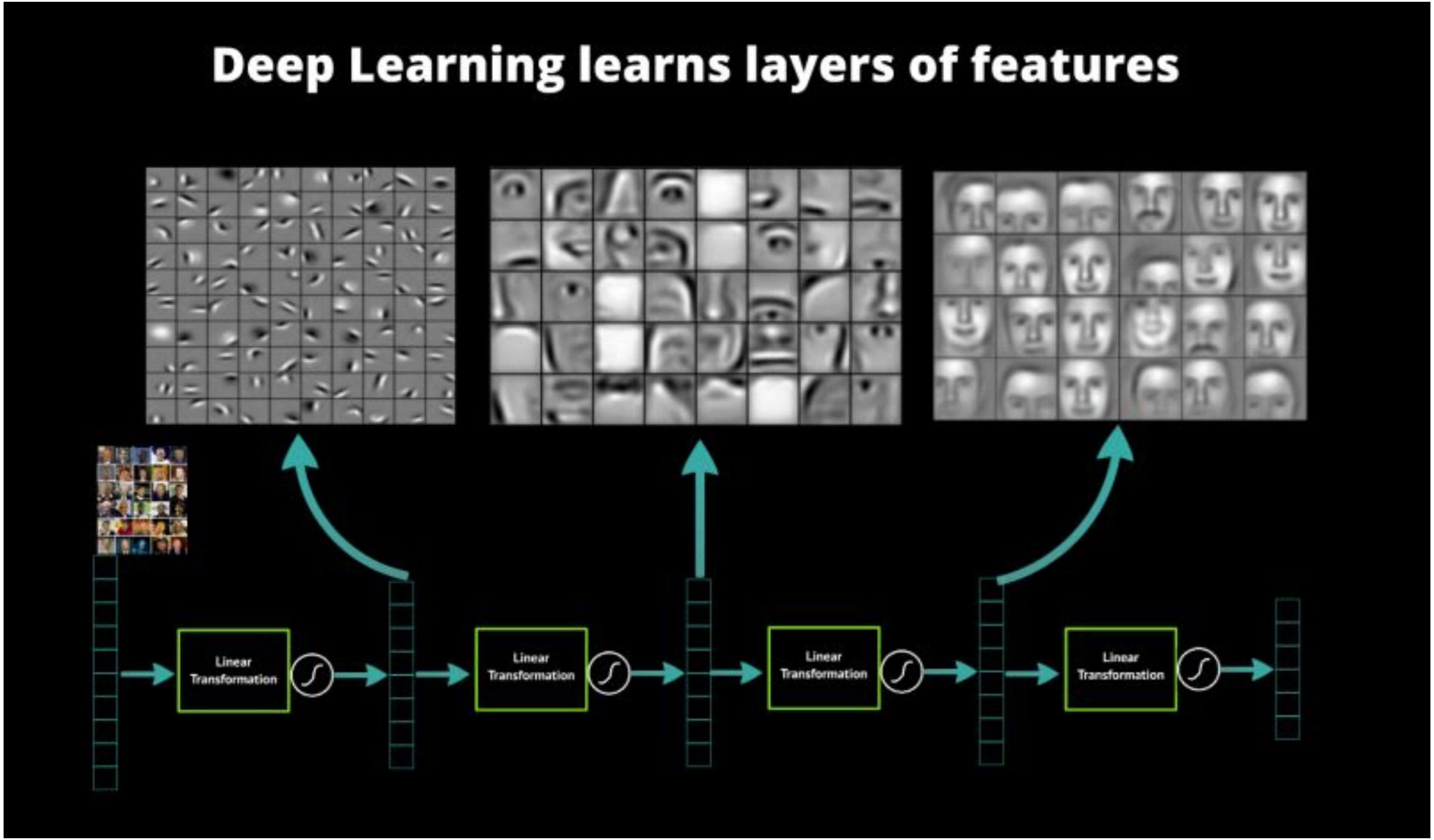
Deep (Feature/Representation) Learning

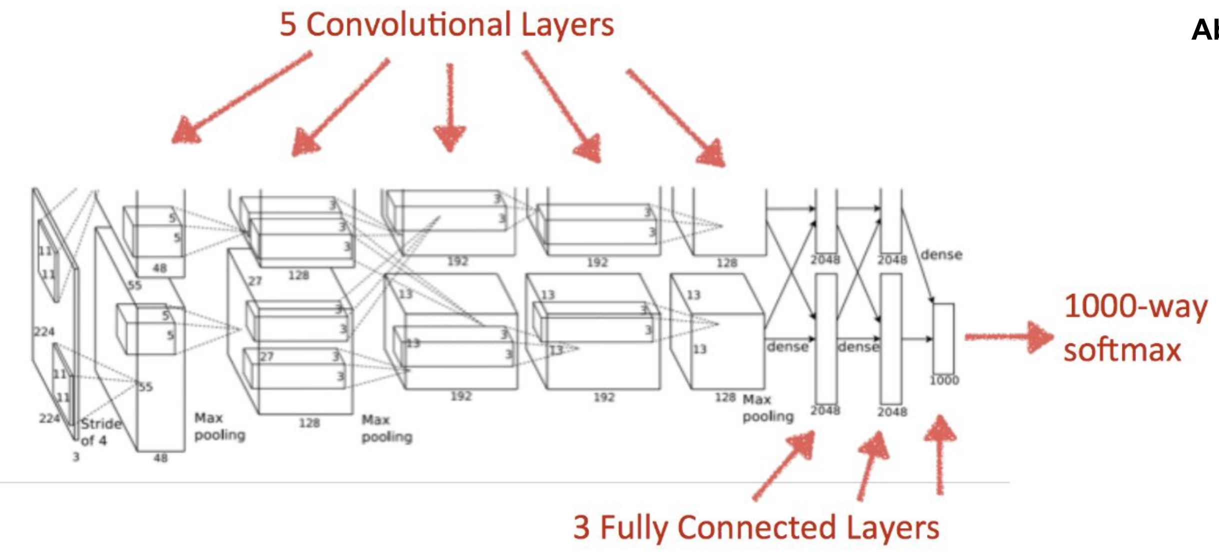
- Move away from feature engineering (still some and some Architechtural design)
- Today learned features generally outperform
- Learn similar gradient based features at early layers

ILSVRC



Li Fei-Fei: ImageNet Large Scale Visual Recognition Challenge, 2014 http://image-net.org/


Deep Learning learns layers of features
Linear Transformation
Linear Transformation
Linear Transformation
Linear Transformation

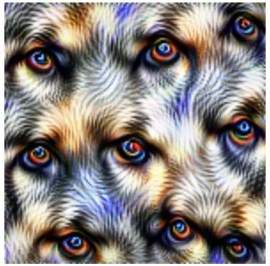


About

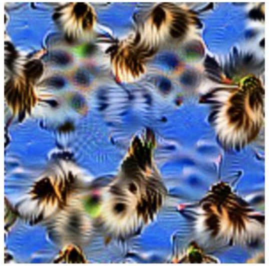
- Yann LeCun. LeNet: http://yann.lecun.com/exdb/lenet/ (1989-1998)
- Inspired by Visual Cortex in cats (receptive fields)
- Designed with image recognition in mind--input and layers often shown as 2D or 3D which may look odd coming from 1D.
- Composition of layers. Rightmost feature layers are most similar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

Layer 4a


Bookshelves


Dog eyes

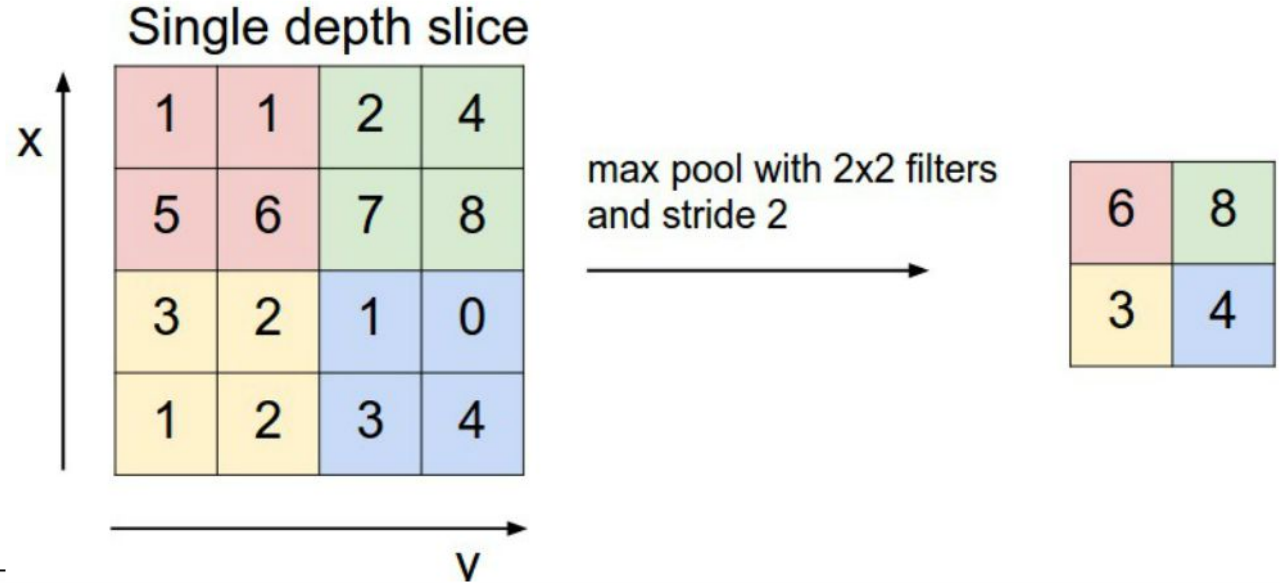

Text, rivets


Birds

In this layer, which follows a pooling step, we see a signficant increase in complexity. We begin to see more complex patterns, and even parts of objects.

Understanding Deep Networks

- Feature visualization: https://distill.pub/2017/feature-visualization/
- Fear and Loathing in LV: https://www.youtube.com/watch?v=oyxSerkkP4o



Visualization of the filter on the image

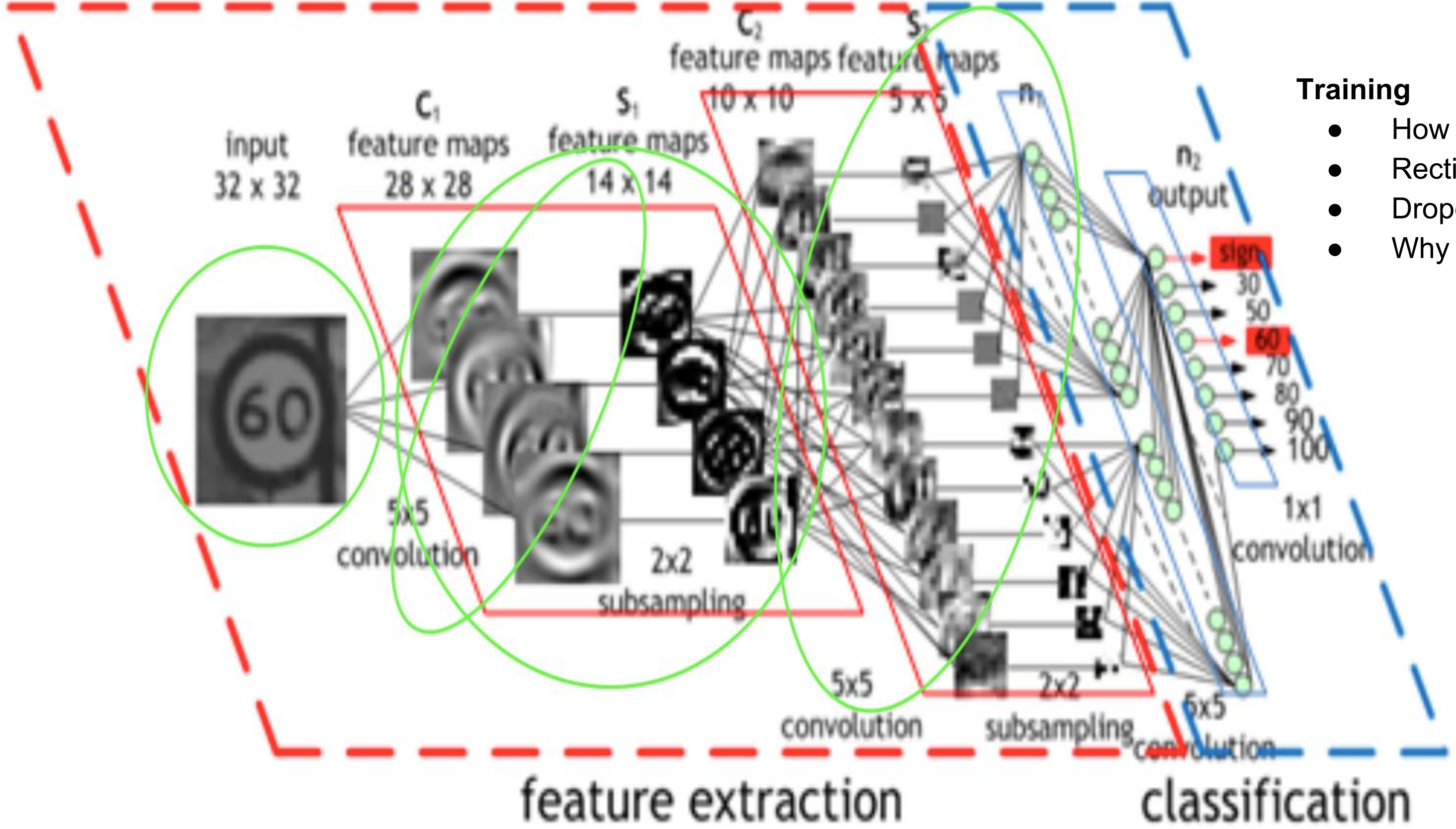
0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

Pixel representation of receptive field

$*$

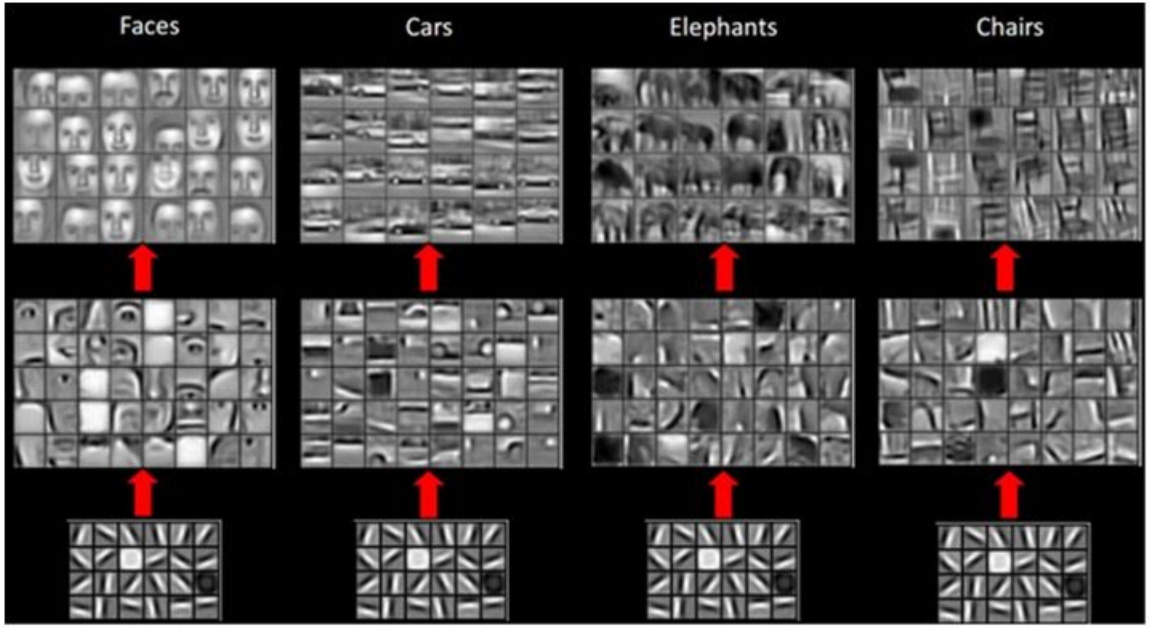
0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter



Training

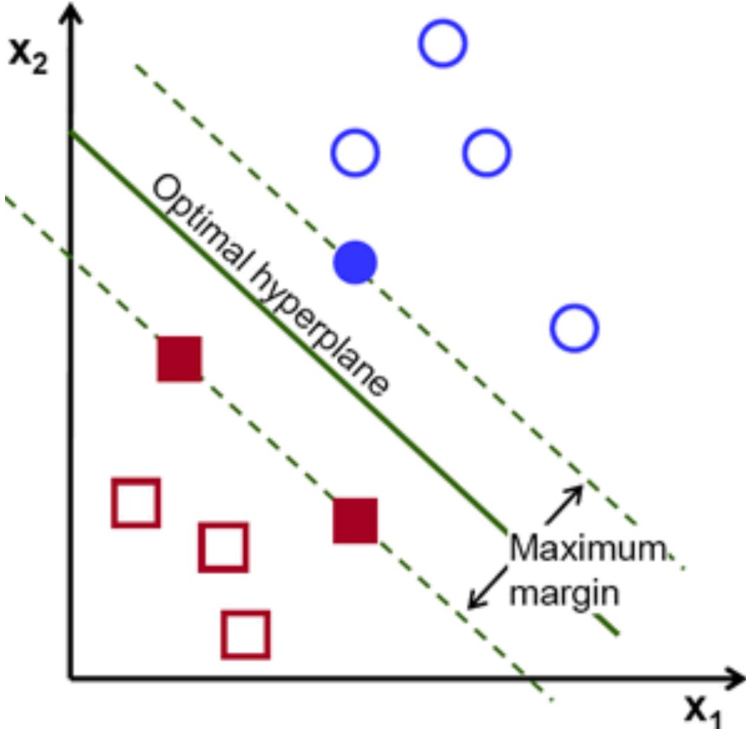
- How many parameters to learn?
- Rectifier activation
- Dropouts
- Why do you think this works well?



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

SVM



Implementations

- LIBSVM/Liblinear -- National Taiwan; used in SK_Learm, e1071, Matlab
- SVMLight/SVMPerf -- Cornell

People

- Vapnik (AT&T, FB)
- Yann LeCun (AT&T, FB)
- Yoshua Bengio (AT&T, Montreal)
- Leon Bottou (AT&T, Google)
- Christopher Bishop (Edinburg, MS)
- Chris Burgess (AT&T, MS)
- Patrick Haffner (AT&T)

Review

- http://www.tristanfletcher.co.uk/SVM%20Explained.pdf
- http://svmlight.joachims.org/

Final Thoughts?