Computer Vision using Convolutional NN

Columbia Water Center
EARTH INSTITUTE | COLUMBIA UNIVERSITY

Oct 2017

Juan Pablo Campos G

Computer vision:

Computers gaining high level understanding from images and videos.
Some Areas:

- Image classification
- Object Recognition
- Object Detection
- Semantic Segmentation
- Instance Segmentation – automatic delineation of
- 3D reconstruction

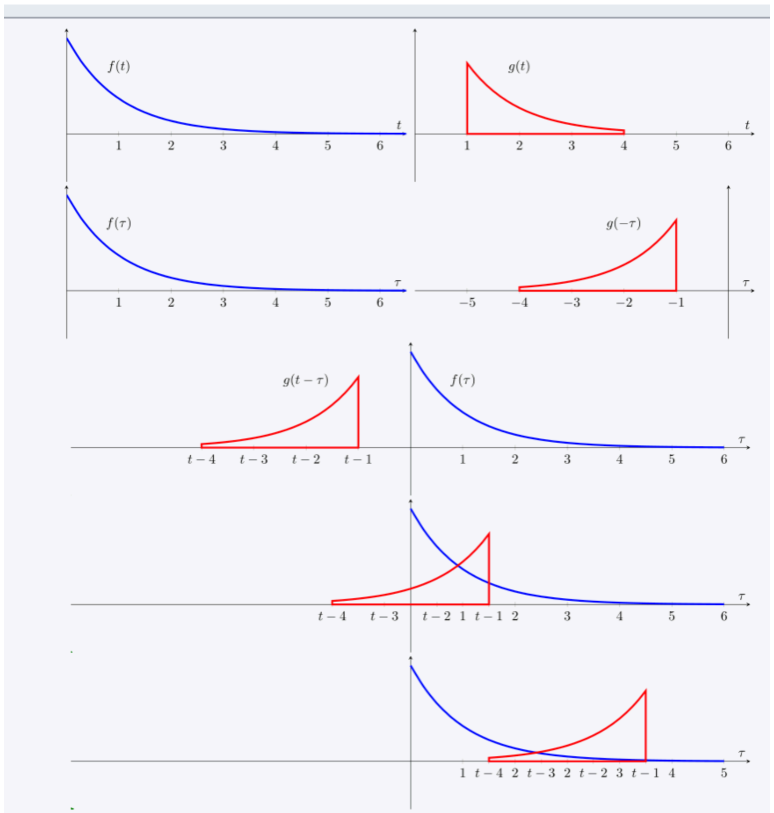
Before ANN:

- SURF, SIFT features + SVMs for object detection
- EigenFaces for Face Recognition
- Kernels for edge detection… Convolutions were used before CNNs!

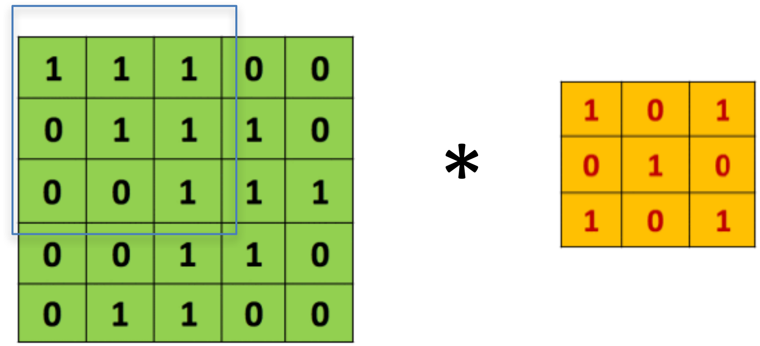
Image Convolutions:

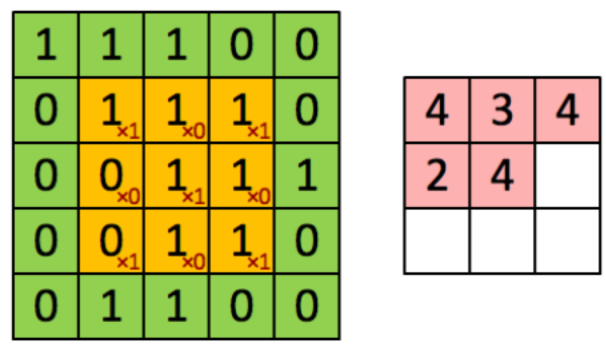
Convolution Mathematics

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n-m]$$



Convolution images

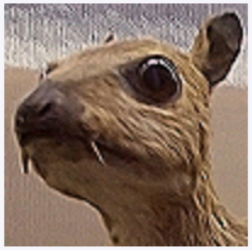




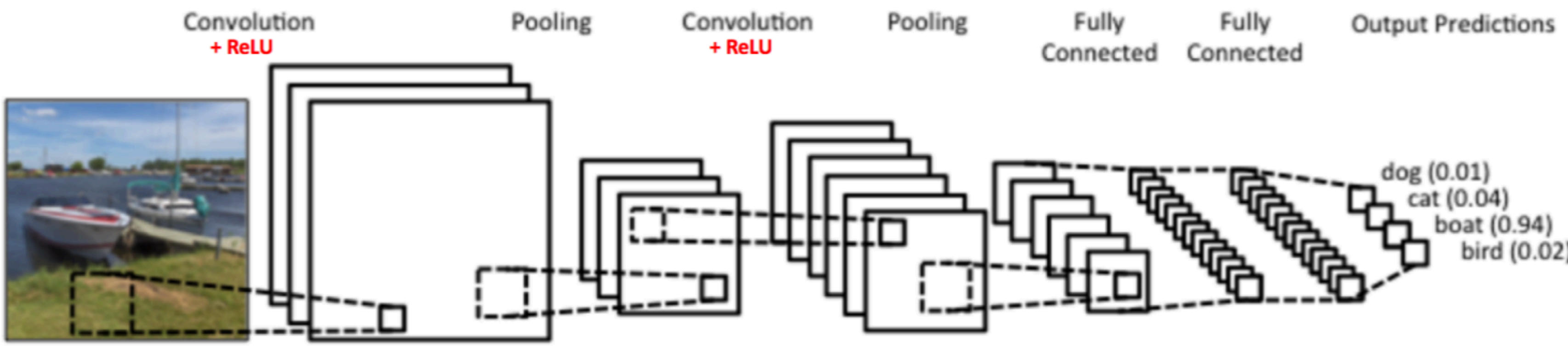
Image

Convolved Feature

Kernels

Edge detection	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (normalized)	$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	

CNN Visualization



Hyperparameters:

Architecture of the Network:

- Number and type of layers
- Depth: number of kernels for a particular layer
- Type of Activation Functions

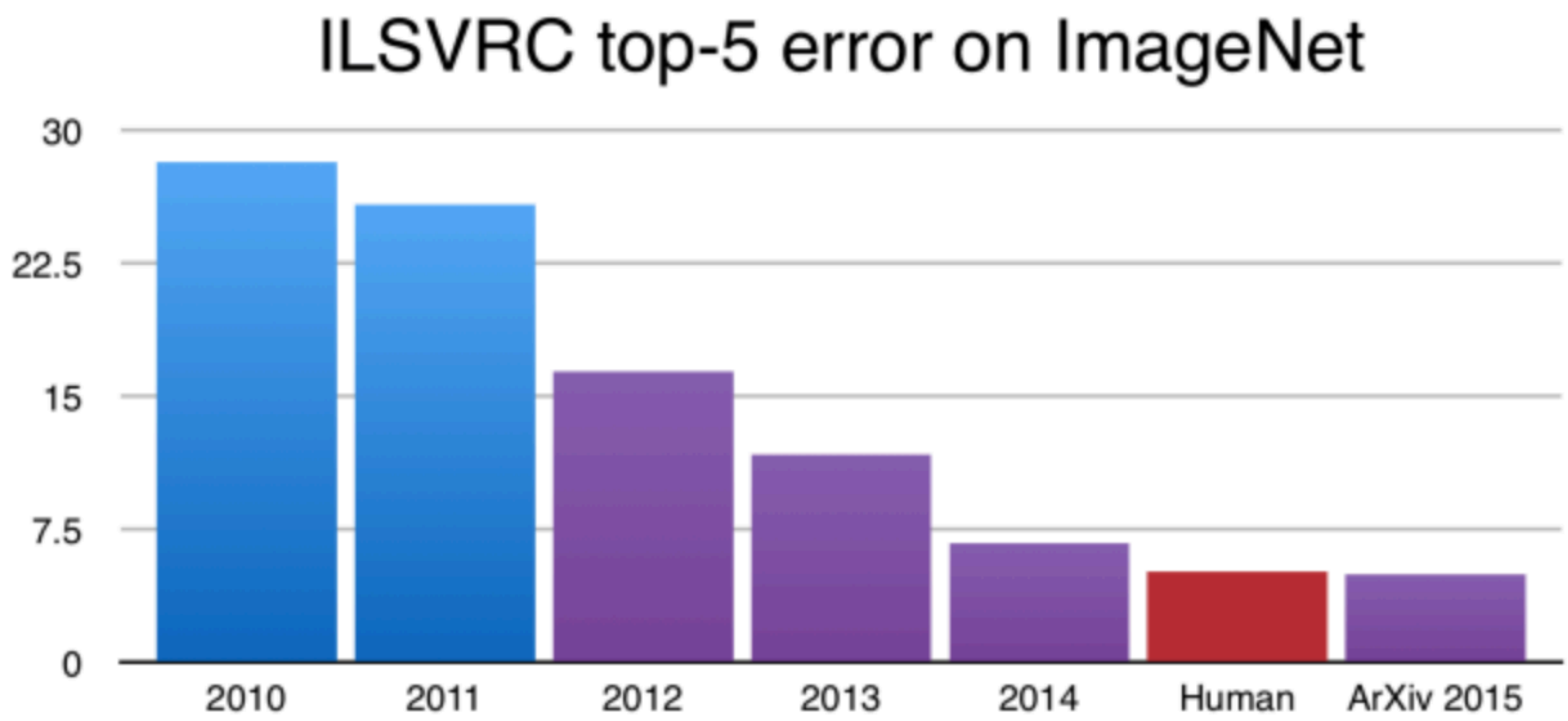
Optimization:

- Learning Rate
- Parameters depending on the function used.
- Batches

History of CNNs

- **LeNet 1990s**
- **1990s to 2012:** convolutional neural network were in incubation.
- **AlexNet (2012) –** In 2012, Alex Krizhevsky (and others) released AlexNet which was a deeper and much wider version of the LeNet and won by a large margin the difficult ImageNet LSVR Challenge in 2012.
- **ZF Net (2013) –** The ILSVRC 2013 winner was a Convolutional Network from Matthew Zeiler and Rob Fergus. It was an improvement on AlexNet by tweaking the architecture hyperparameters.
- **GoogLeNet (2014) –** The ILSVRC 2014 winner was a Convolutional Network from Szegedy et al.from Google. Its main contribution was the development of an *Inception Module* that dramatically reduced the number of parameters in the network (4M, compared to AlexNet with 60M).
- **VGGNet (2014) –** The runner-up in ILSVRC 2014 was the network that became known as the VGGNet. Its main contribution was in showing that the depth of the network (number of layers) is a critical component for good performance.
- **ResNets (2015) –** Residual Network developed by Kaiming He (and others) was the winner of ILSVRC 2015.
- **DenseNet (August 2016) –** Recently published by Gao Huang (and others), the Densely Connected Convolutional Network has each layer directly connected to every other layer in a feed-forward fashion. The DenseNet has been shown to obtain significant improvements over previous state-of-the-art architectures on five highly competitive object recognition benchmark tasks. Check out the Torch implementation here.

ImageNet model errors



CNNs for Image Detection and Object Segmentation

Extending the achievements of Alexnet to Instance Segmentation.

- R-CNN
- Fast R-CNN
- Faster R-CNN
- Mask R-CNN

Ideas for next session

Implementation and challenges:

- Optimizers: which optimizers to use?
- Hyper parameters, how to choose them?
- Model overfitting:
    - Regularization
    - Visualize plots: learning curve, etc.
- Which functions to use for the middle layers?
- Which functions to use in the last layers?