Convolutional neural networks
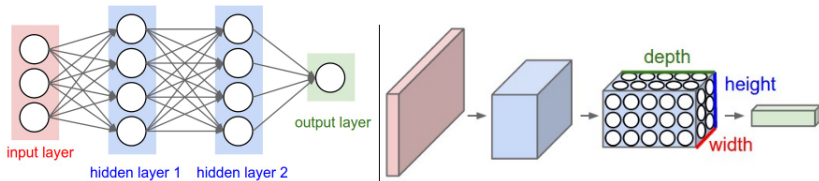
Mikhail Yurushkin
m.yurushkin@broutonlab.com

April 2, 2019
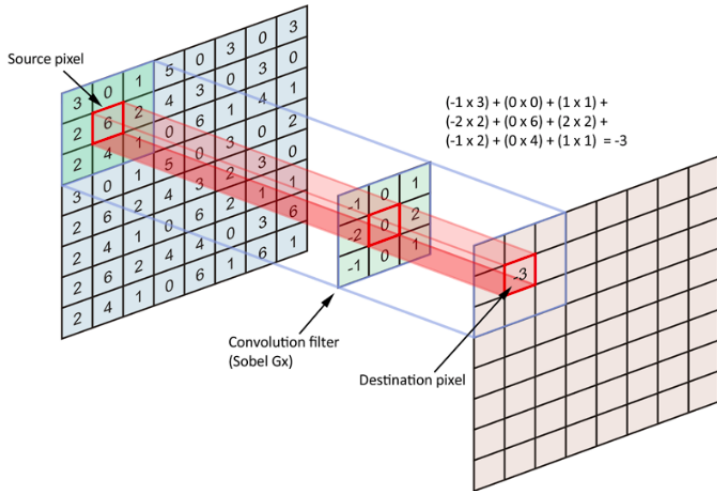
Brouton **Lab**

https://broutonlab.com

Regular neural networks and ConvNets

- ▶ Regular Neural Nets don't scale well to full images.
- ▶ ConvNet have neurons arranged in 3 dimensions: width, height, depth

Convolutional Layer

Source pixel

Convolution filter
(Sobel Gx)

Destination pixel

$(-1 \times 3) + (0 \times 0) + (1 \times 1) +$
$(-2 \times 2) + (0 \times 6) + (2 \times 2) +$
$(-1 \times 2) + (0 \times 4) + (1 \times 1) = -3$

Parameters

- ▶ Kernel Size - is a filter applied in slidding window.
- ▶ Stride - controls how the filter convolves around the input volume
- ▶ Depth - controls number of neurons in a layer that connect to the same region of the input volume
- ▶ Padding - controls the output volume spatial size

Spatial Pooling layers (Max/Average)
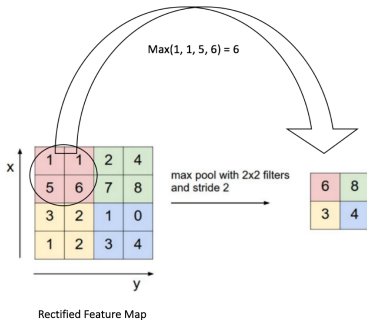
- Reduces the dimensionality of each feature map
- Spatial Pooling can be of different types: Max, Average, Sum etc.
- Seems to be will be discarded in the future (see Striving for Simplicity: The All Convolutional Net'14).
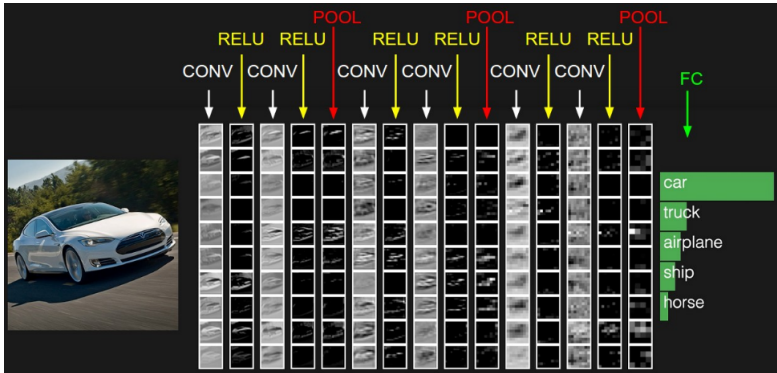
LeNet (1990s)
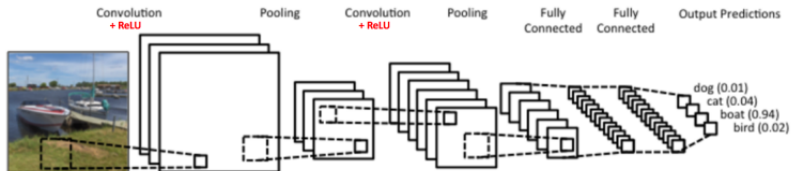
1. Very first convolutional neural network
2. Classifies images of 10 classes (dog, cat, bird etc)
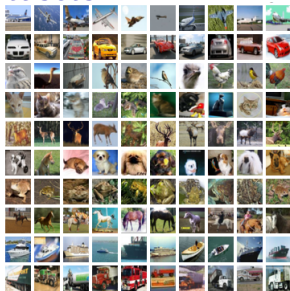
Comparison of state of art architectures
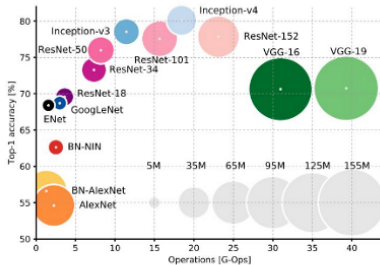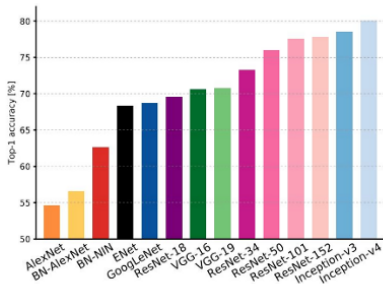
An Analysis of Deep Neural Network Models for Practical Applications, 2017.

Trainable filters

Benchmarks and datasets

- ▶ PASCAL VOC, COCO, ImageNet, CIFAR (2D classification, object detection)

- ▶ KITTI Vision Benchmark (tereo, optical flow, visual odometry, 3D object detection and 3D tracking)

- ▶ CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.