

State-of-the-art recognition systems are based on *convolutional* neural networks

Public databases of face images serve as benchmarks:

Labeled Faces in the Wild (LFW, <http://vis-www.cs.umass.edu/lfw>)
> 13,000 images of celebrities, 5,749 different identities

YouTube Faces Database (YTF, <http://www.cs.tau.ac.il/~wolf/ytfaces>)
3,425 videos, 1,595 different identities

Private face image datasets:

(Facebook) Social Face Classification dataset
4.4 million face photos, 4,030 different identities

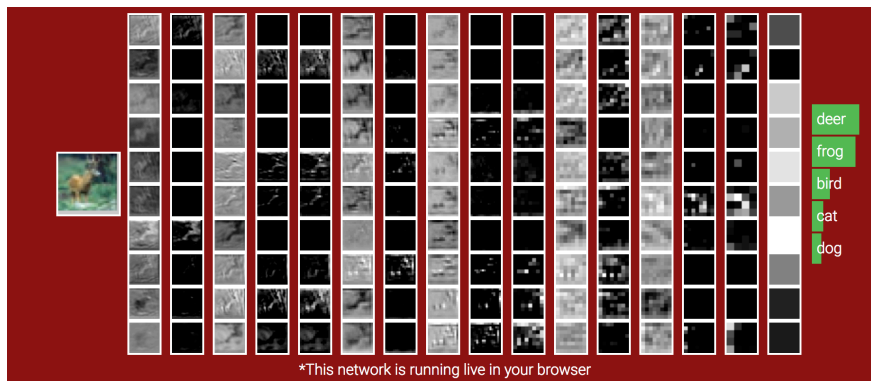
(Google) 100-200 million face images, ~ 8 million different identities

	LFW	YTF
Facebook DeepFace	97.4%	91.4%
Google FaceNet	99.6%	95.1%
Human performance	97.5%	89.7%



Convolutional Neural Networks (CNNs)

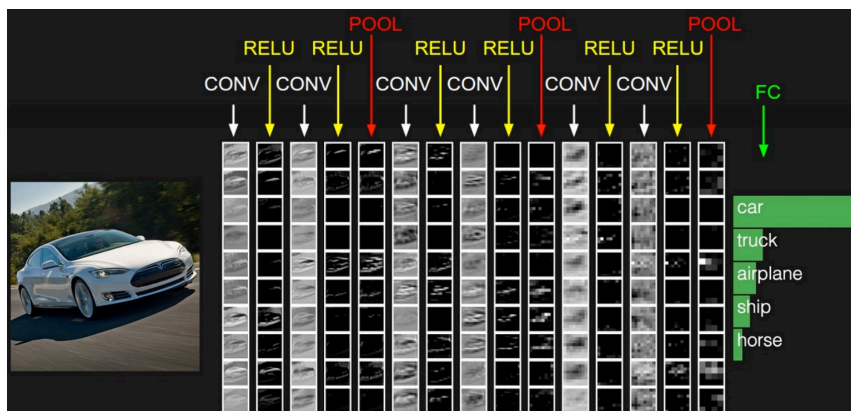
Fei-Fei Li, Justin Johnson, Serena Yeung (<http://cs231n.stanford.edu/>)



*This network is running live in your browser

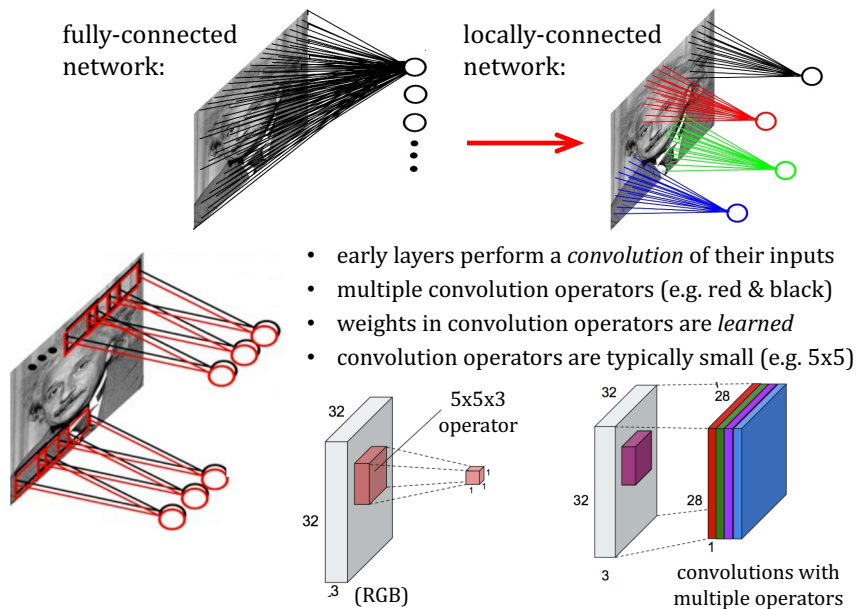
The Convolutional Neural Network in this example is classifying images live in your browser using Javascript, at about 10 milliseconds per image. It takes an input image and transforms it through a series of functions into class probabilities at the end. The transformed representations in this visualization can be loosely thought of as the activations of the neurons along the way. The parameters of this function are learned with backpropagation on a dataset of (image, label) pairs. This particular network is classifying **CIFAR-10** images into one of 10 classes and was trained with **ConvNetJS**. Its exact architecture is [conv-relu-conv-relu-pool|x3-fc-softmax, for a total of 17 layers and 7000 parameters. It uses 3x3 convolutions and 2x2 pooling regions. By the end of the class, you will know exactly what all these numbers mean.

Sample stages of a CNN

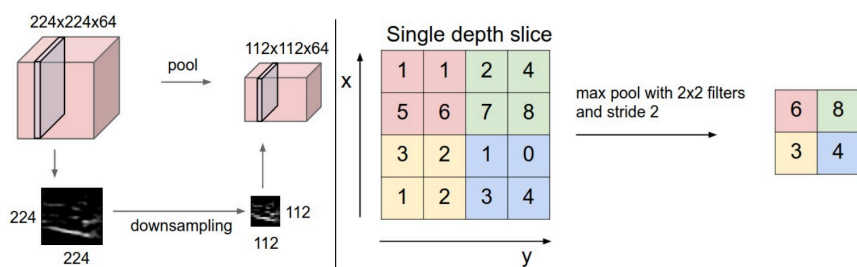
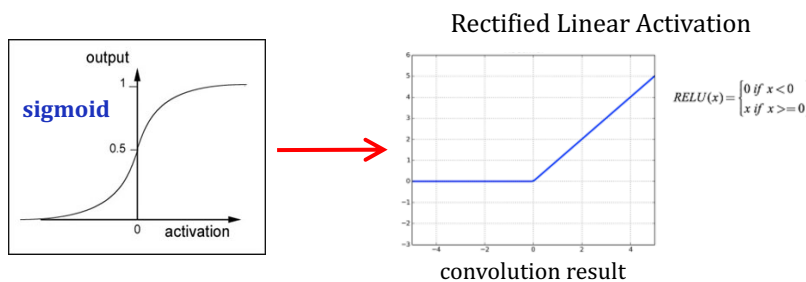


CONV: "convolution" layer with weights that are learned
 RELU: "rectified linear unit" applies an activation function
 POOL: "pooling" selects maximum value in small neighborhoods
 FC: "fully-connected" neural network

Convolutional layer

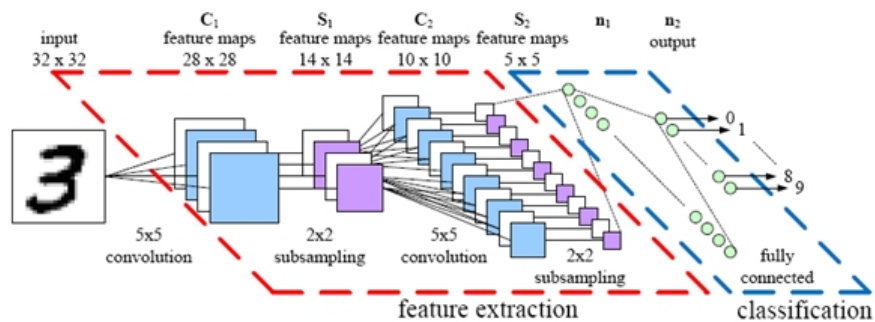


ReLU & max pooling layers



Adding a fully-connected neural net layer

Recognizing digits from the MNIST database with a CNN:



LeCun, Bottou, Bengio, Haffner (1998)