A Crash Course in Computer Vision with Keras

February 2019

hi, I'm Helen!

- Machine Learning Engineer
- Sidewalk Toronto Fellow
- Toronto Women's Data Group
- Mathematics, coffee, poetry

DESSA

mathemakitten.github.io

mathemakitten

www.linkedin.com/in/helen-ngo/

fast.ai

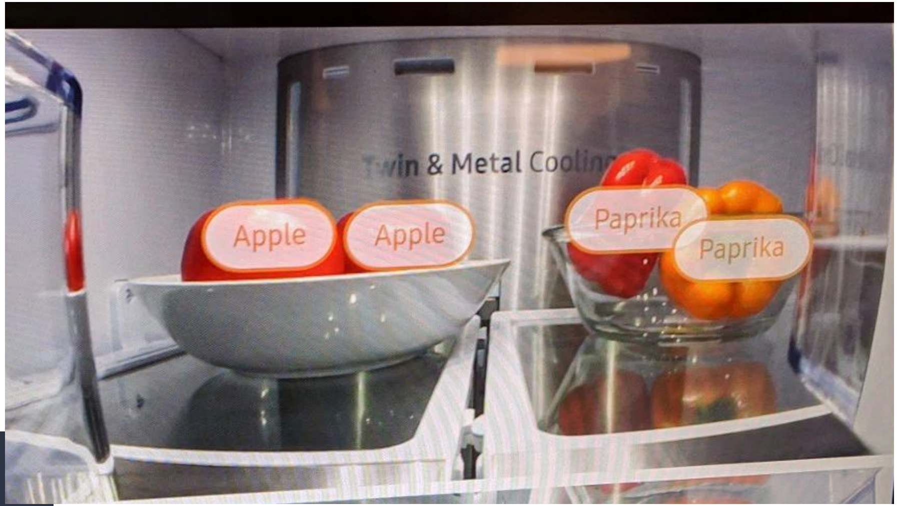
Practical Deep Learning for Coders

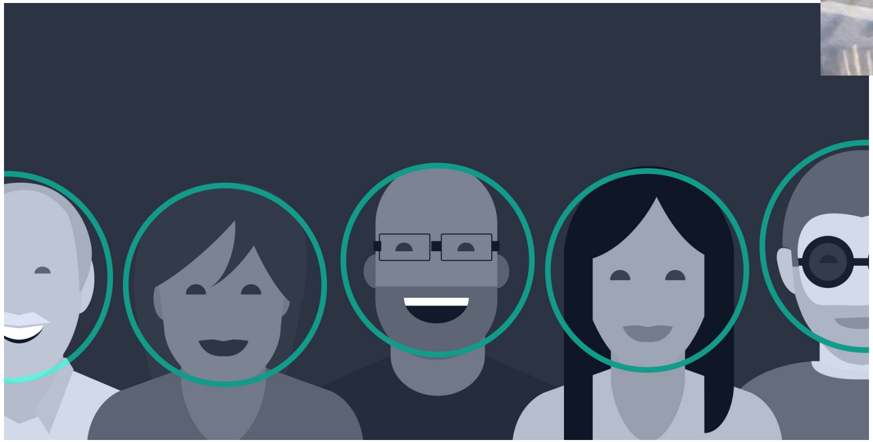
(2017 edition)

Jeremy Howard (Enlitic) & Rachel Thomas (USF)

1 year of coding + high school math

Computer vision is a subfield of **deep learning** which extracts understanding from digital images or video.


Twin & Metal Cooling
Apple
Apple
Paprika
Paprika



François Chollet @fchollet · Jan 31

Using Keras to automate supernovas identification, potentially cutting in half the time it would take astronomers to discover supernovas:

How 3 engineers built a record-breaking supernova identification sys...

Pop into Dessa's offices and you'll soon find traces of the company's fascination with outer space. A Lego replica of Saturn V, the rocket that m...

medium.com

4 89 279

Geological formation, formation

(geology) the geological features of the earth

1808 pictures

86.24% Popularity Percentile

Wordnet IDs

Numbers in brackets: (the number of synsets in the subtree).

- ImageNet 2011 Fall Release (32326)
  - plant, flora, plant life (4486)
  - geological formation, formation (1
    - aquifer (0)
    - beach (1)
    - cave (3)
    - cliff, drop, drop-off (2)
    - delta (0)
    - diapir (0)
    - folium (0)
    - foreshore (0)
    - ice mass (10)
    - lakefront (0)
    - massif (0)
    - monocline (0)
    - mouth (0)
    - natural depression, depression (
    - natural elevation, elevation (41
    - oceanfront (0)
    - range, mountain range, range of
    - relict (0)
    - ridge, ridgeline (2)
    - ridge (0)
    - shore (7)
    - slope, incline, side (17)
    - spring, fountain, outflow, outpo
    - talus, scree (0)
    - vein, mineral vein (1)
    - volcanic crater, crater (2)
    - wall (0)

Treemap Visualization | Images of the Synset | Downloads

ImageNet 2011 Fall Release › Geological formation, formation

Research Prediction Competition

ImageNet Object Localization Challenge

Identify the objects in images

ImageNet · 33 teams · 11 years to go

Overview

Competition Description

While It's pretty easy for people to identify subtle differences in photos, computers still have a ways to go. Visually similar items are tough for computers to count, like this overlapping bunch of bananas

ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

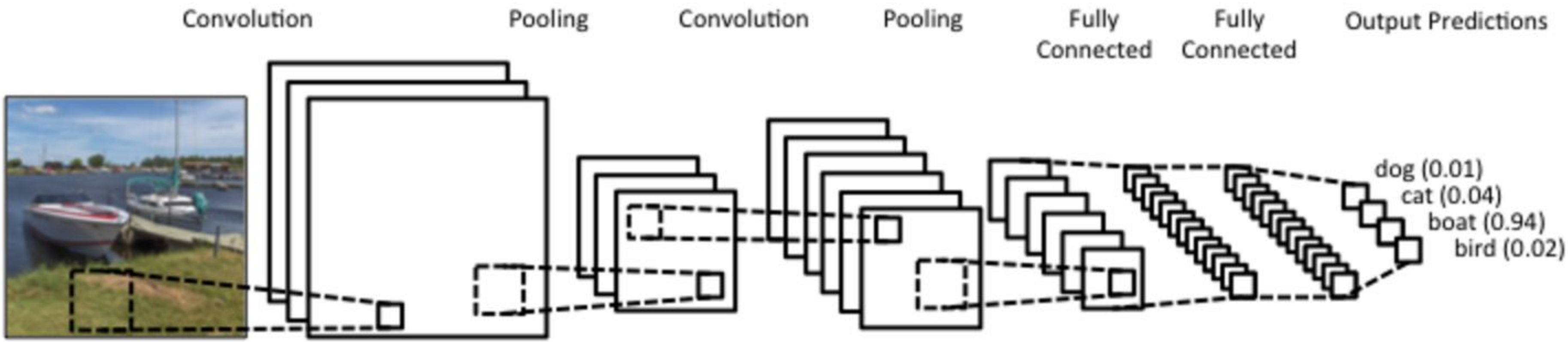
Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

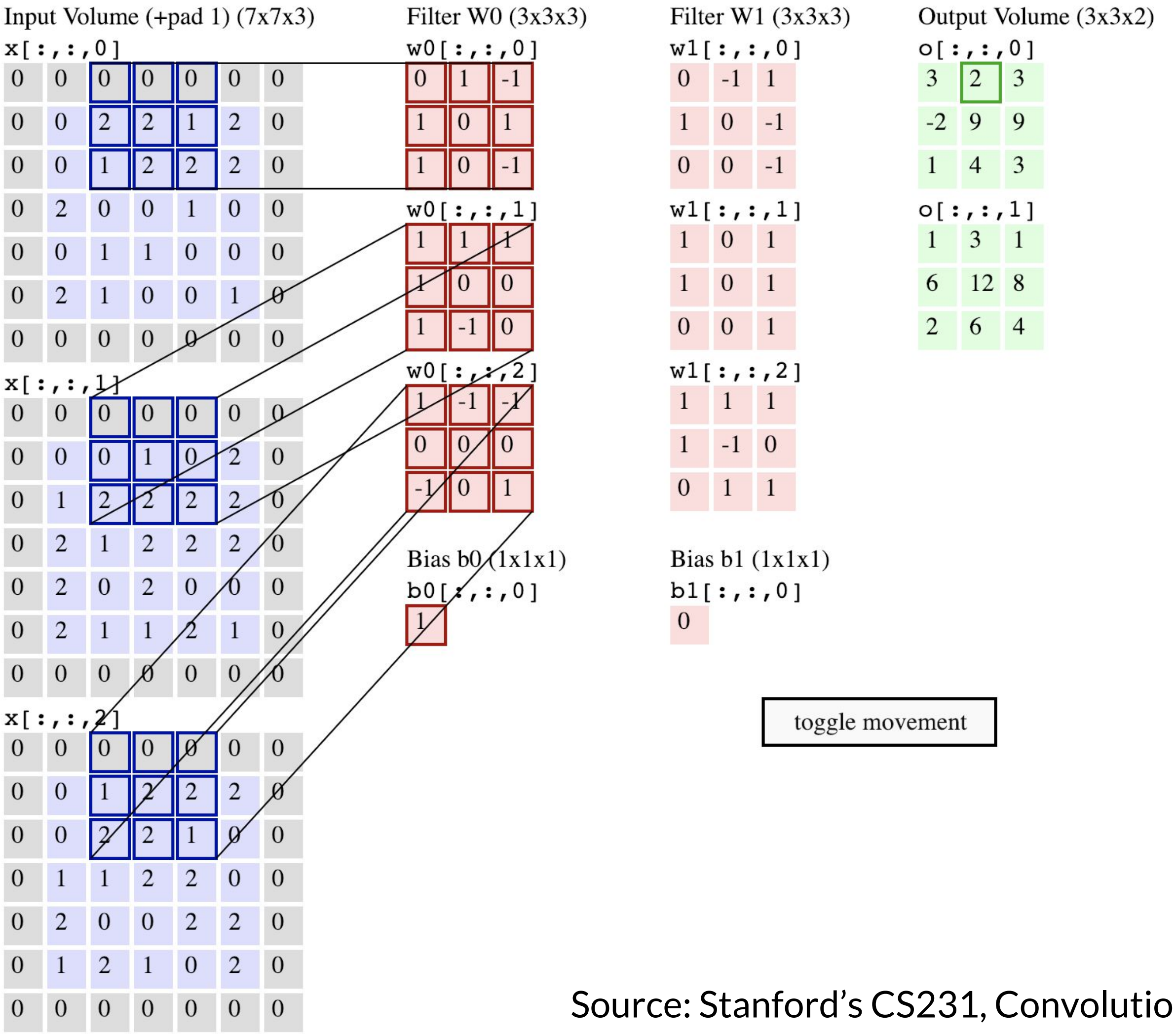
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Abstract

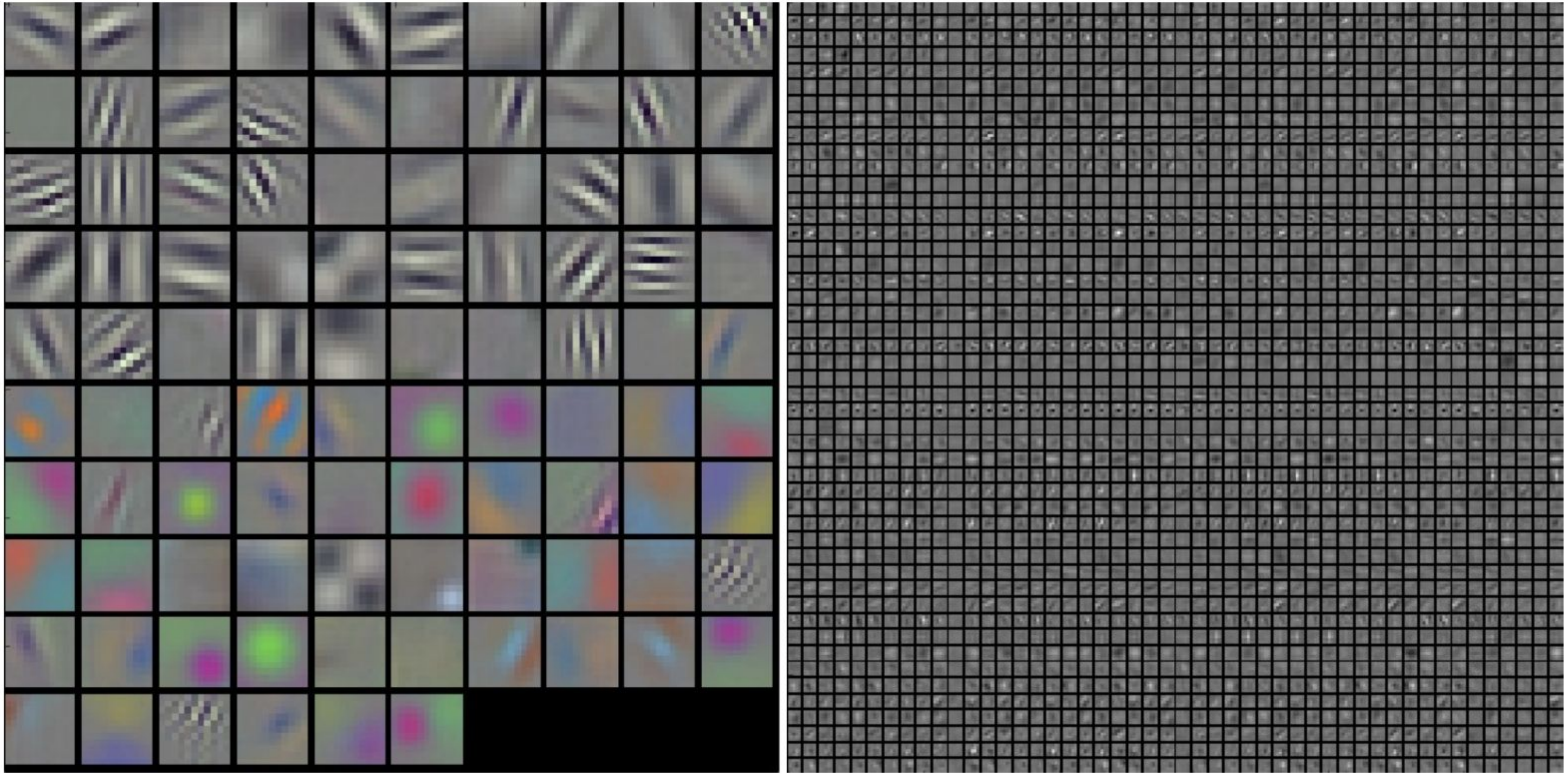
We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called "dropout" that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.

anatomy of a convnet



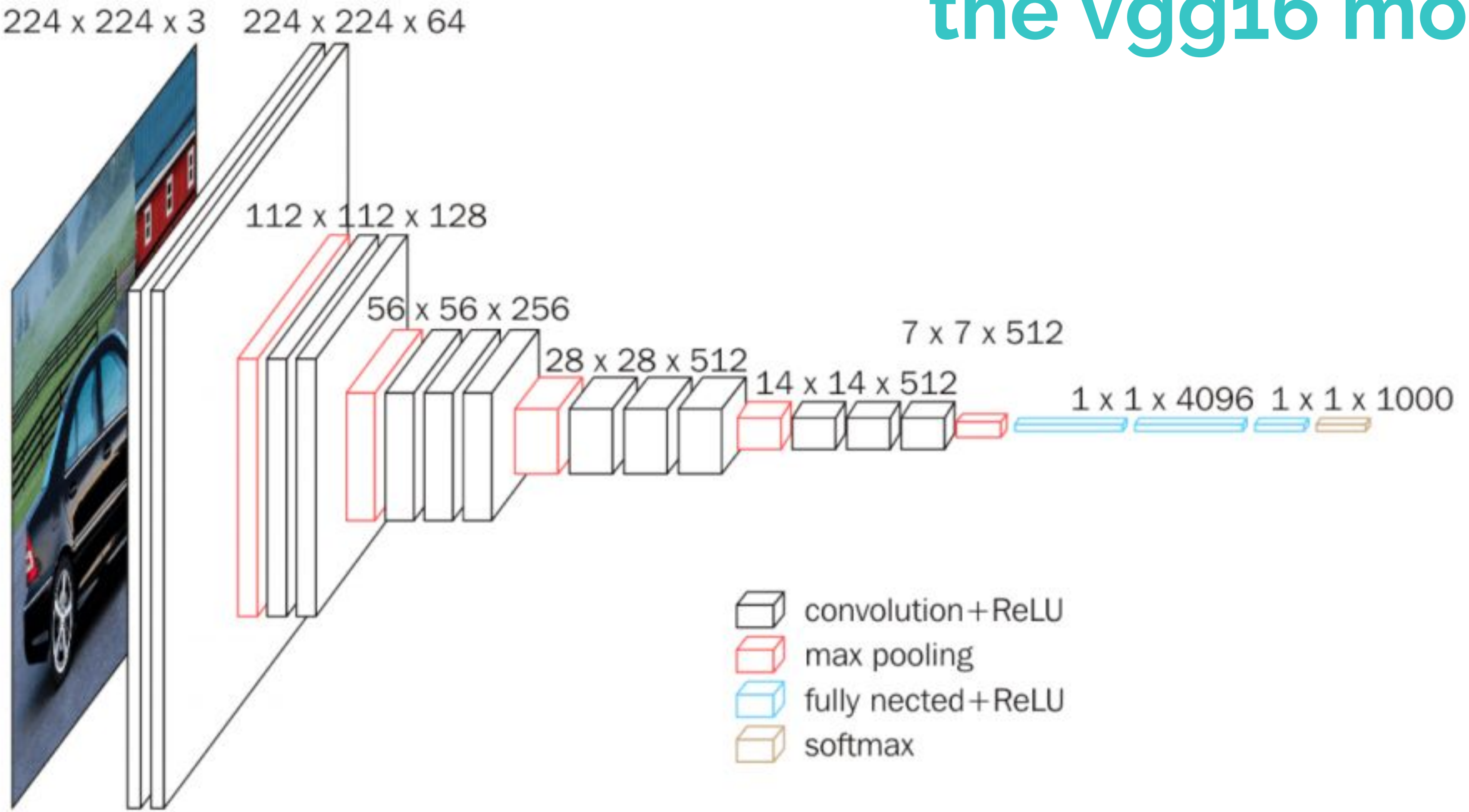


Source: Stanford's CS231, Convolutional Neural Networks



Typical-looking filters on the first CONV layer (left), and the 2nd CONV layer (right) of a trained AlexNet. Notice that the first-layer weights are very nice and smooth, indicating nicely converged network. The color/grayscale features are clustered because the AlexNet contains two separate streams of processing, and an apparent consequence of this architecture is that one stream develops high-frequency grayscale features and the other low-frequency color features. The 2nd CONV layer weights are not as interpretable, but it is apparent that they are still smooth, well-formed, and absent of noisy patterns.

the vgg16 model



https://neurohive.io/en/popular-networks/vgg16/

Colaboratory by Google

- NVIDIA Tesla K80 GPUs
  - 12 GB of RAM
- 12 hours of runtime
- TPU access!

Workshop!

Open with Colaboratory:

github.com/mathemakitten/fastai-colab/blob/master/devhub_workshop.ipynb