convolutional neural networks, swift and iOS 12

by brett koonce

june 7th, 2018

overview

- goal: give you a boost along ml path
- do cnn/nn review, things to study, discuss a framework for tackling problems
- current tools, different ideas on how to approach things
- state of the art, beyond

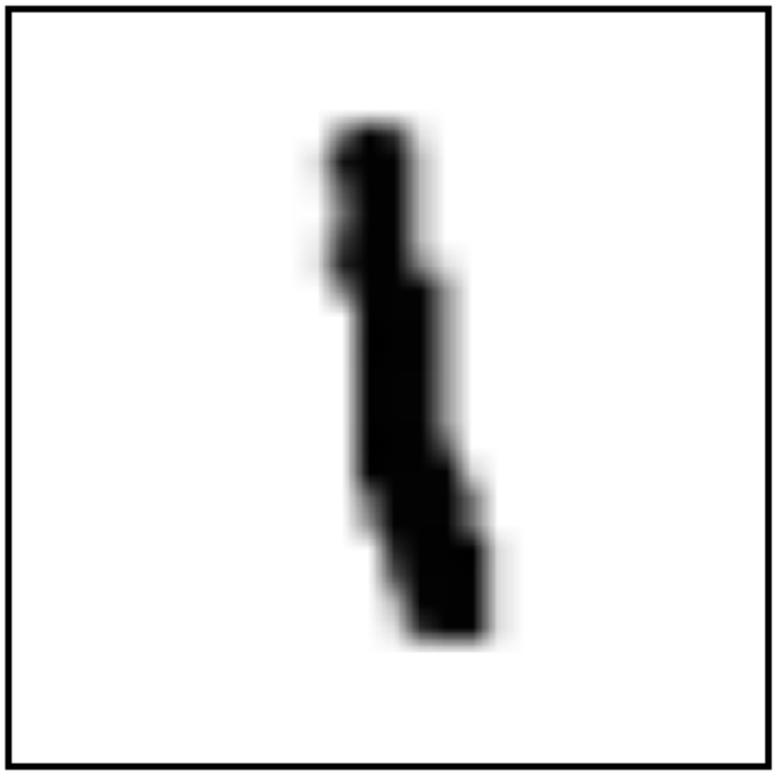
enter the temple

machine learning

ml concepts

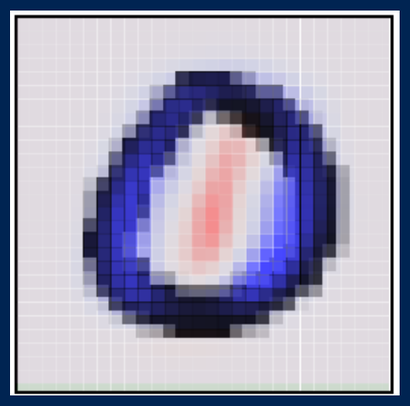
- an input (numbers, image, audio, video)
- known data (supervised learning)
- combine to produce function/black box
- train model, use on unknown data
- goals: quality, size, complexity

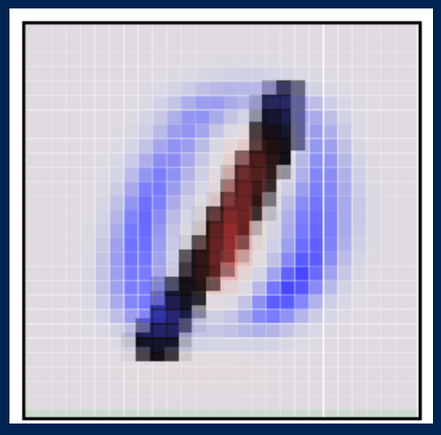
mnist: hello world



≈

0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	.6	.8	0	0	0	0	0	0
0	0	0	0	0	0	.7	1	0	0	0	0	0	0
0	0	0	0	0	0	.7	1	0	0	0	0	0	0
0	0	0	0	0	0	.5	1	.4	0	0	0	0	0
0	0	0	0	0	0	0	1	.4	0	0	0	0	0
0	0	0	0	0	0	0	1	.4	0	0	0	0	0
0	0	0	0	0	0	0	1	.7	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	0	0	0	0
0	0	0	0	0	0	0	.9	1	.1	0	0	0	0
0	0	0	0	0	0	0	.3	1	.1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0



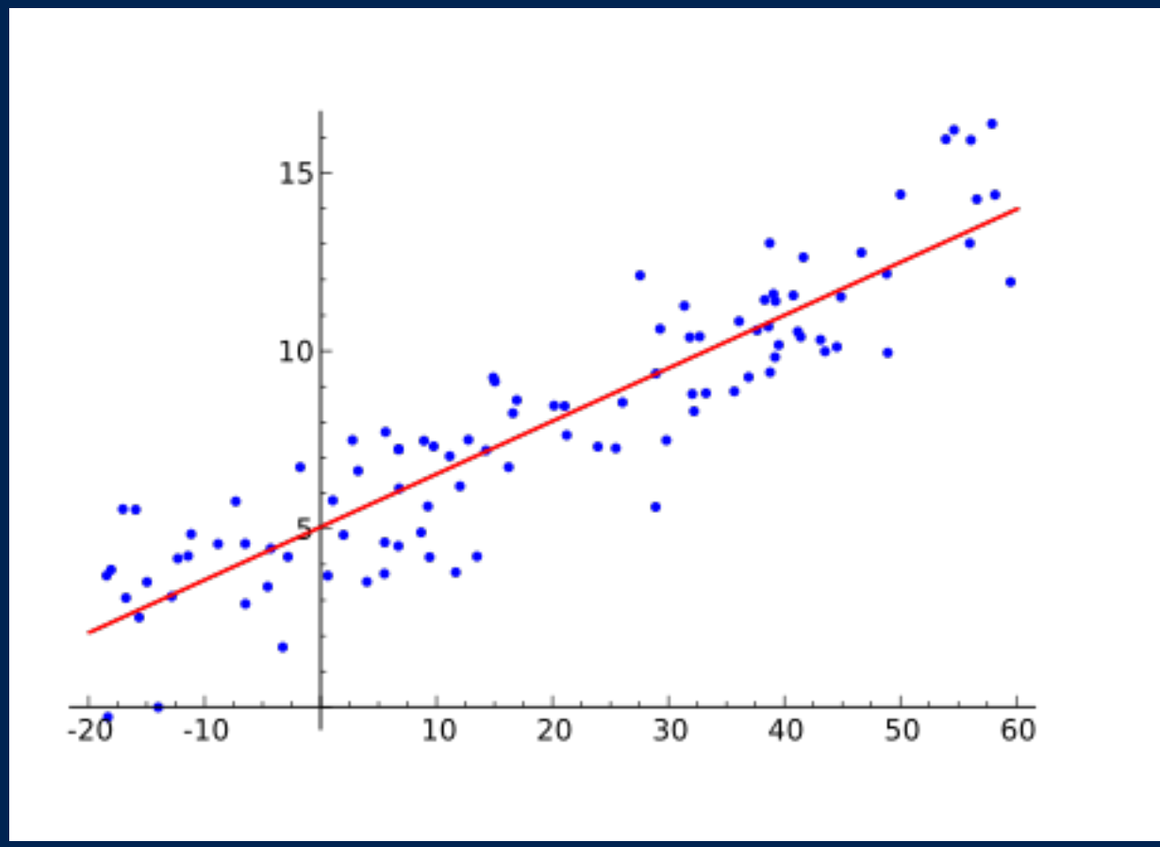


five easy pieces

- **linear regression**
- **random forests**
- **support vector machines**
- **gradient boosted trees**
- **neural network**

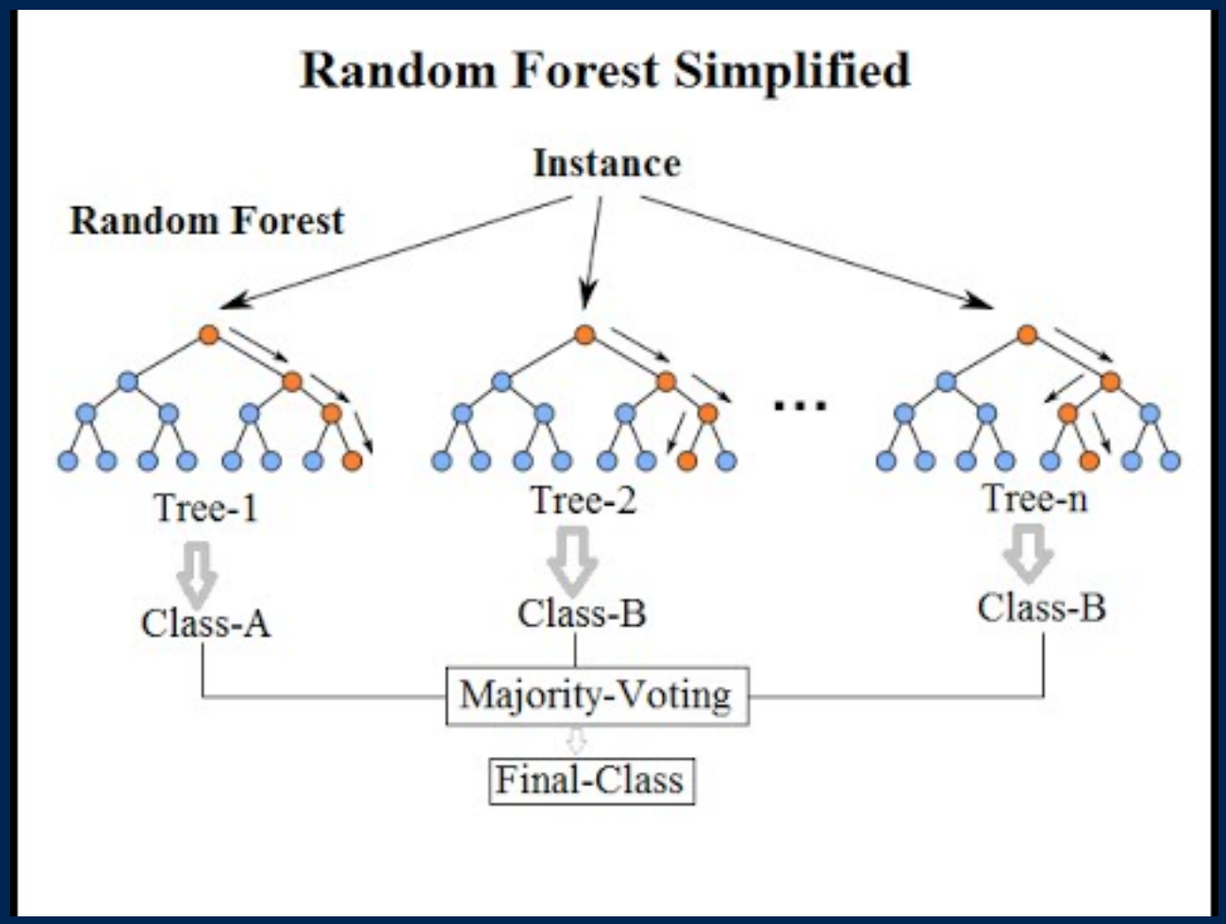
linear regression

- points —> slope == linear regression
- think of adding all the points you have as one epoch



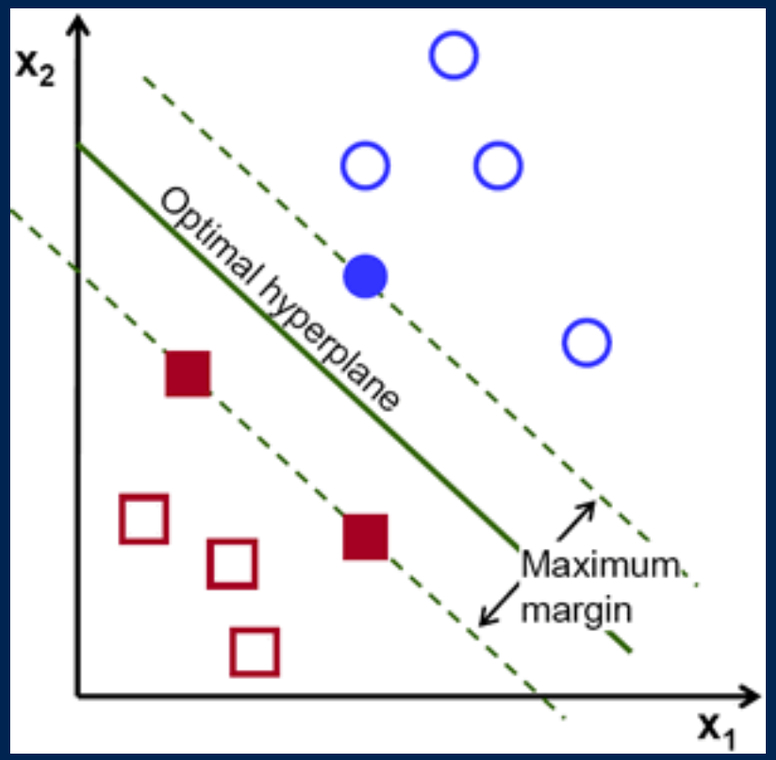
random forests

- **decision trees + sampling —> rf**
- **very robust to noise**
- **cheap**



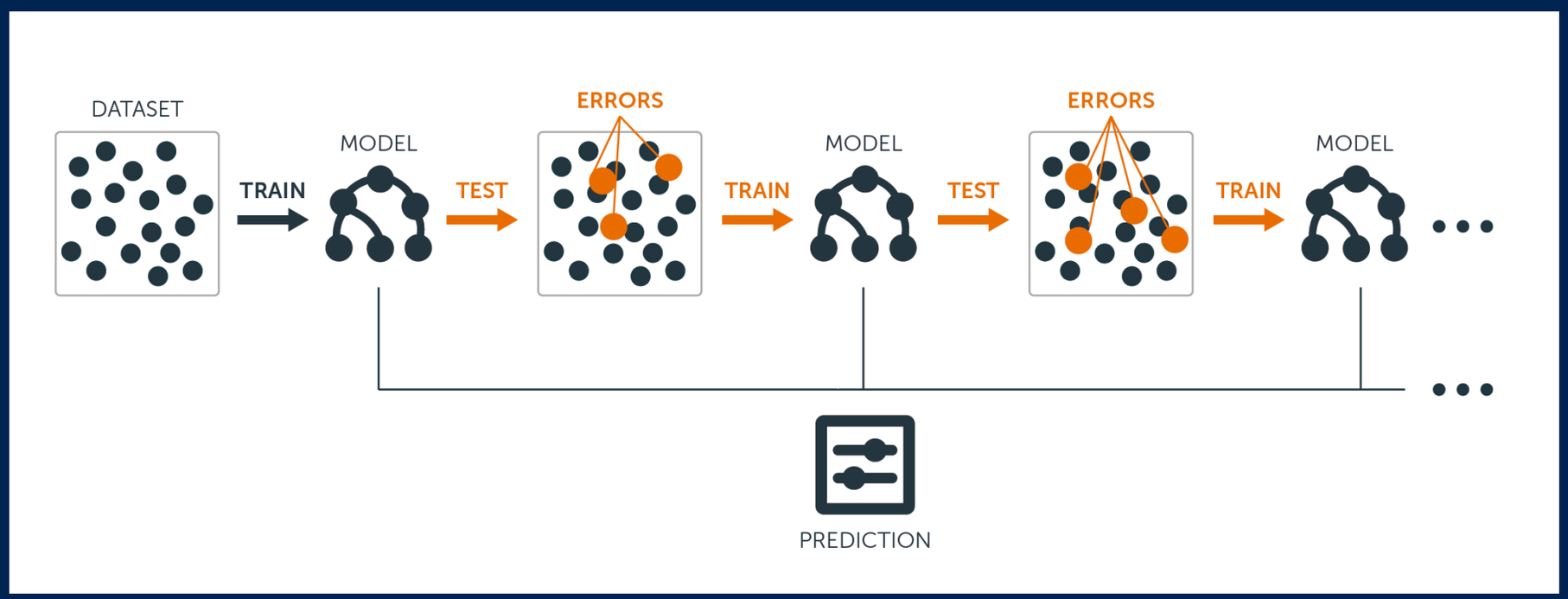
support vector machines

- mathy approach to separating data
- kernel tricks for finding features



gradient boosted trees

- random forests + linear regression == tree ensembles == gbt
- powerful when you have feature data and want to find interesting correlations



neural network

- 1 layer nn == linear regression
- 2 layer nn == neural network
- feed forward neural network
- tensorflow.js mnist layers example

Neural Networks

- Backfed Input Cell
- Input Cell
- Noisy Input Cell
- Hidden Cell
- Probablistic Hidden Cell
- Spiking Hidden Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- Different Memory Cell
- Kernel
- Convolution or Pool

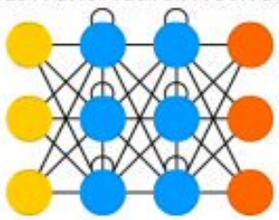
Perceptron (P)

Feed Forward (FF)

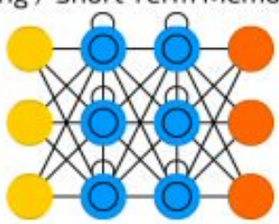
Radial Basis Network (RBF)

Deep Feed Forward (DFF)

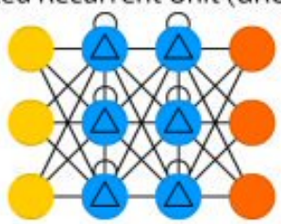
Recurrent Neural Network (RNN)



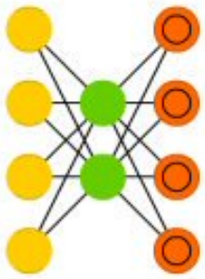
Long / Short Term Memory (LSTM)



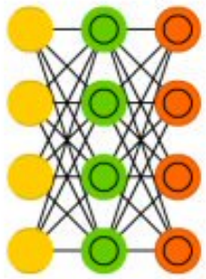
Gated Recurrent Unit (GRU)



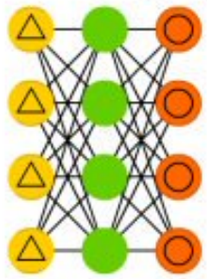
Auto Encoder (AE)



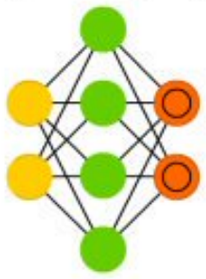
Variational AE (VAE)



Denoising AE (DAE)



Sparse AE (SAE)

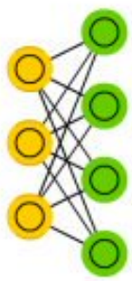


Markov Chain (MC)

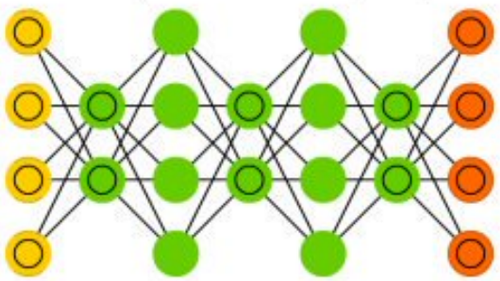
Hopfield Network (HN)

Boltzmann Machine (BM)

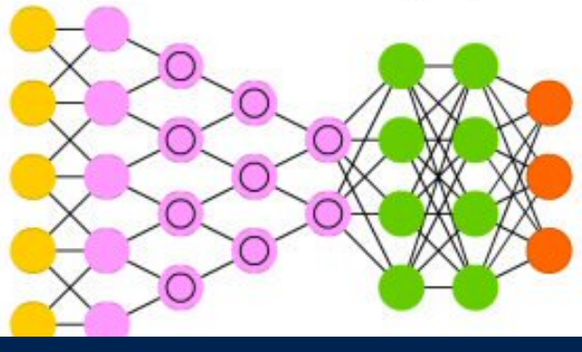
Restricted BM (RBM)



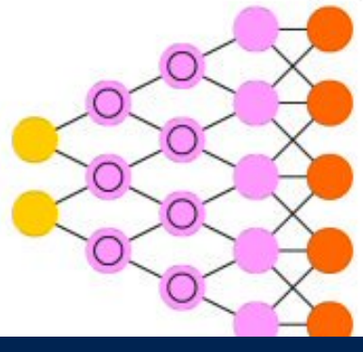
Deep Belief Network (DBN)



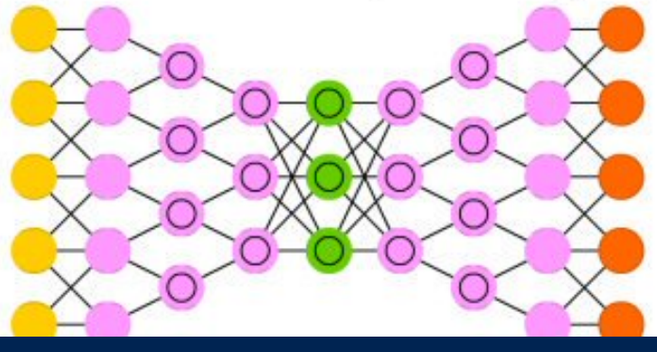
Deep Convolutional Network (DCN)



Deconvolutional Network (DN)



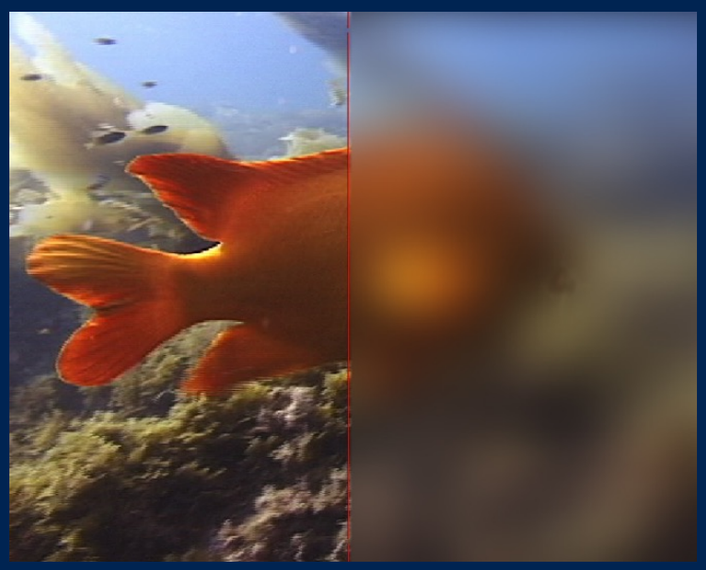
Deep Convolutional Inverse Graphics Network (DCIGN)



convolutions

- **convolution == matrix math == a[x] + b**
- **transform[image] + offset == i1 —> i2**

1/16	1/8	1/16
1/8	1/4	1/8
1/16	1/8	1/16



-1	0	1
-2	0	2
-1	0	1

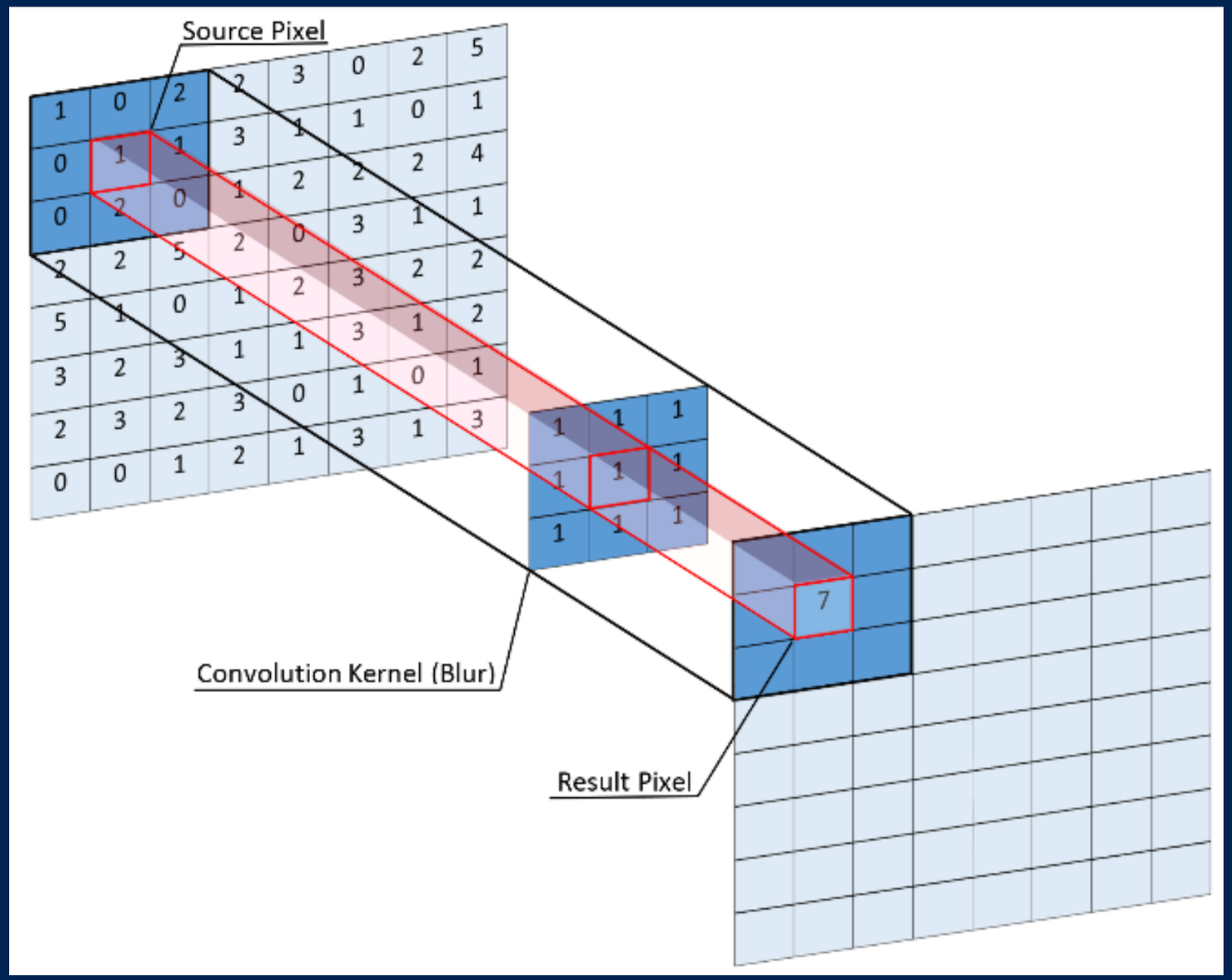
Horizontal

-1	-2	-1
0	0	0
-1	-2	-1

Vertical

convolutions v2

- **size transforms up**
- **size transforms down**

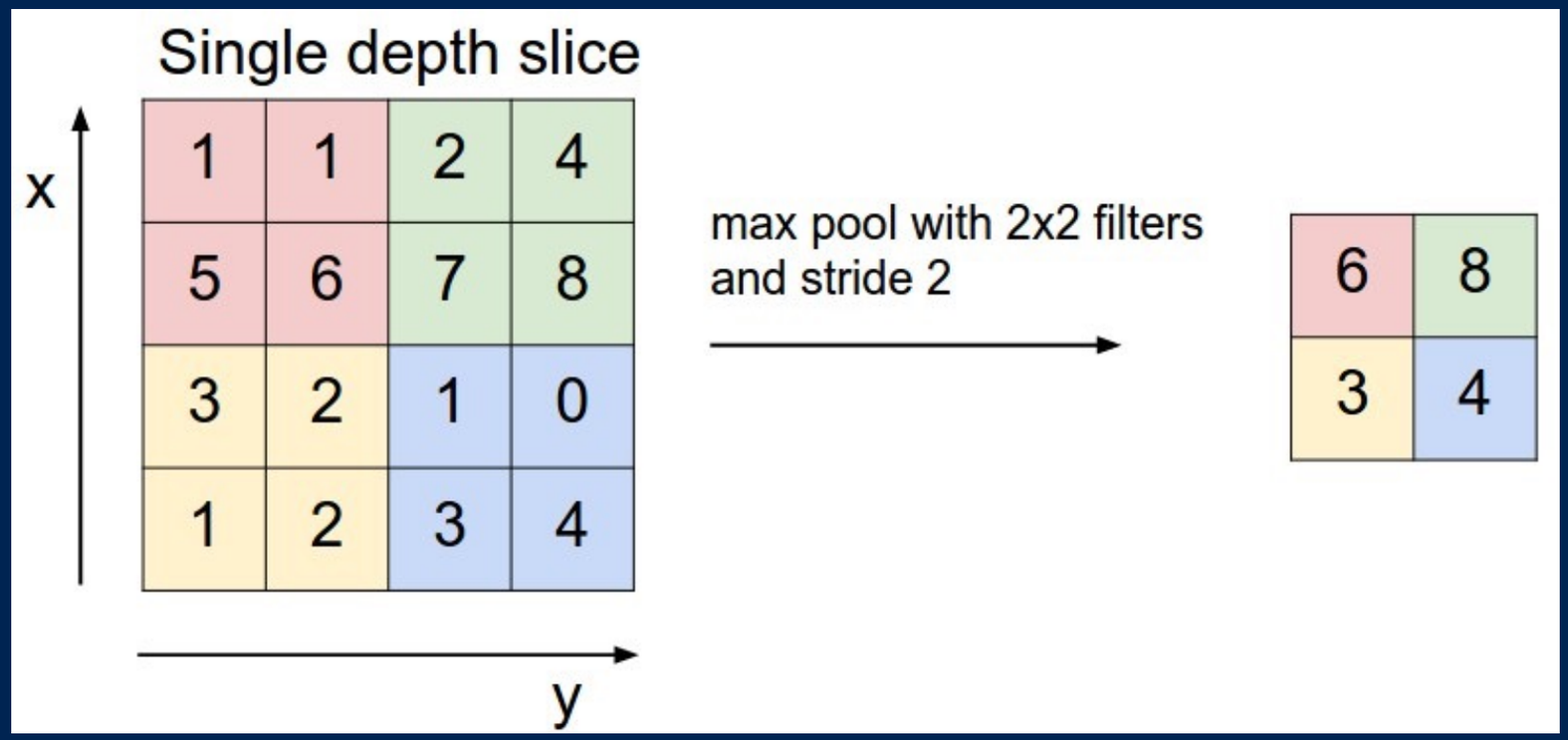


conv: 3x3 striding

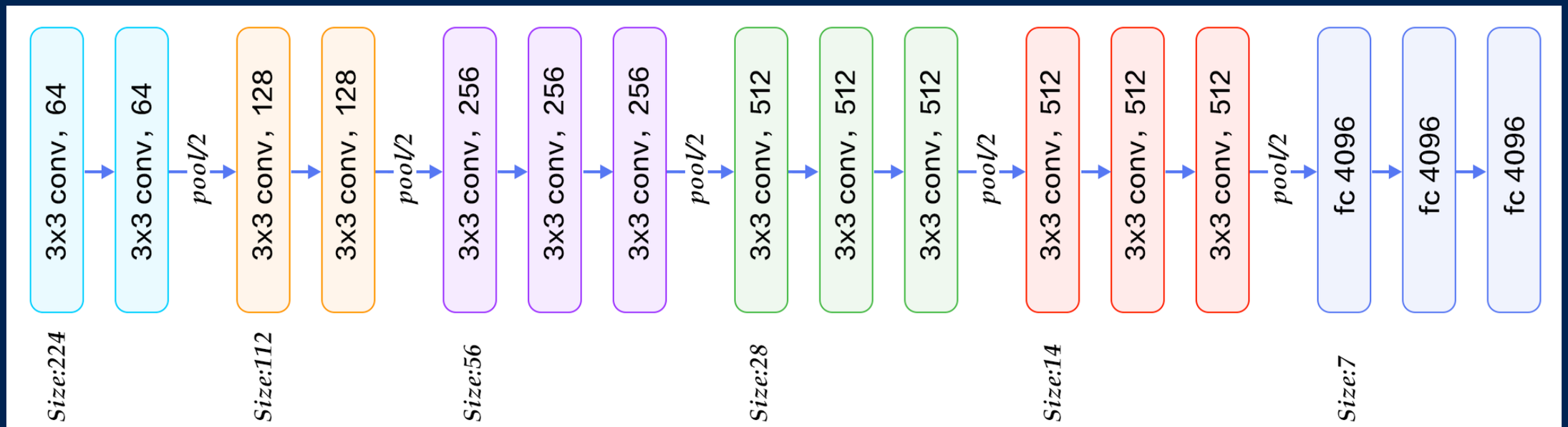
- upsample input image into chunks
- cs231n.github.io
- deepfish

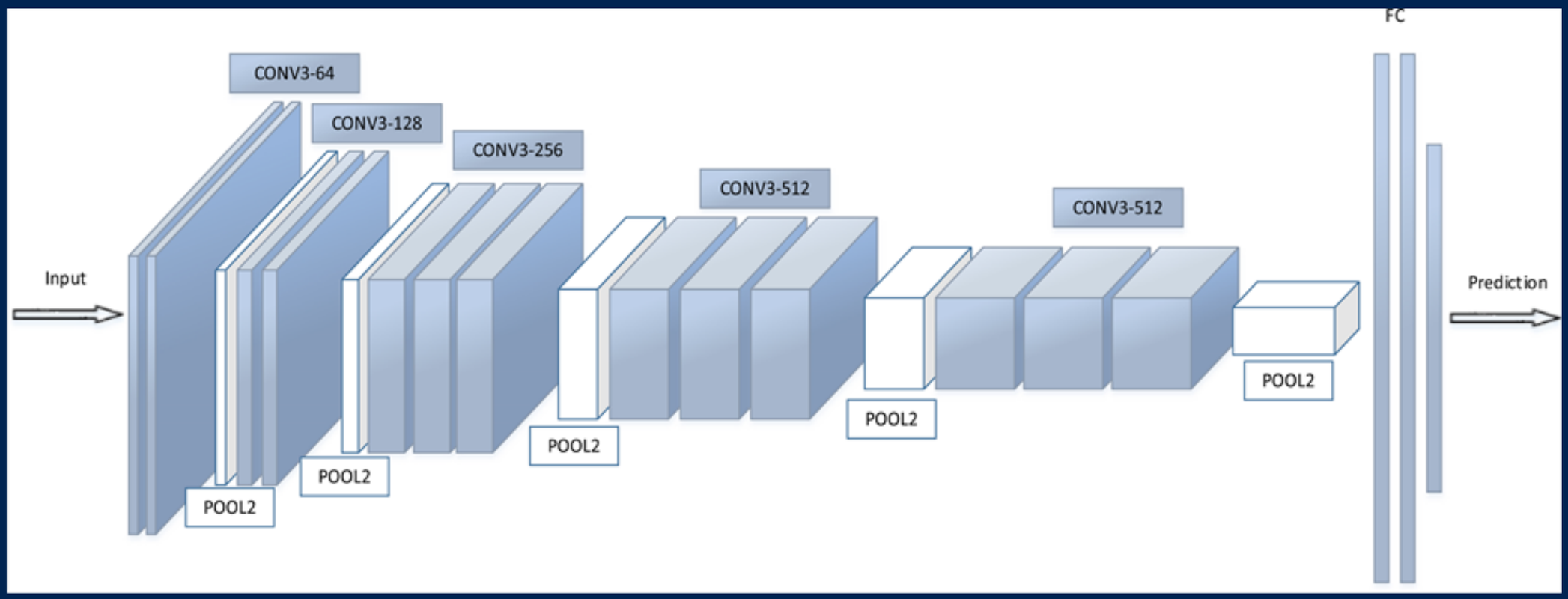
conv: maxpool

- striding produces a lot of samples
- next: downsample the data!



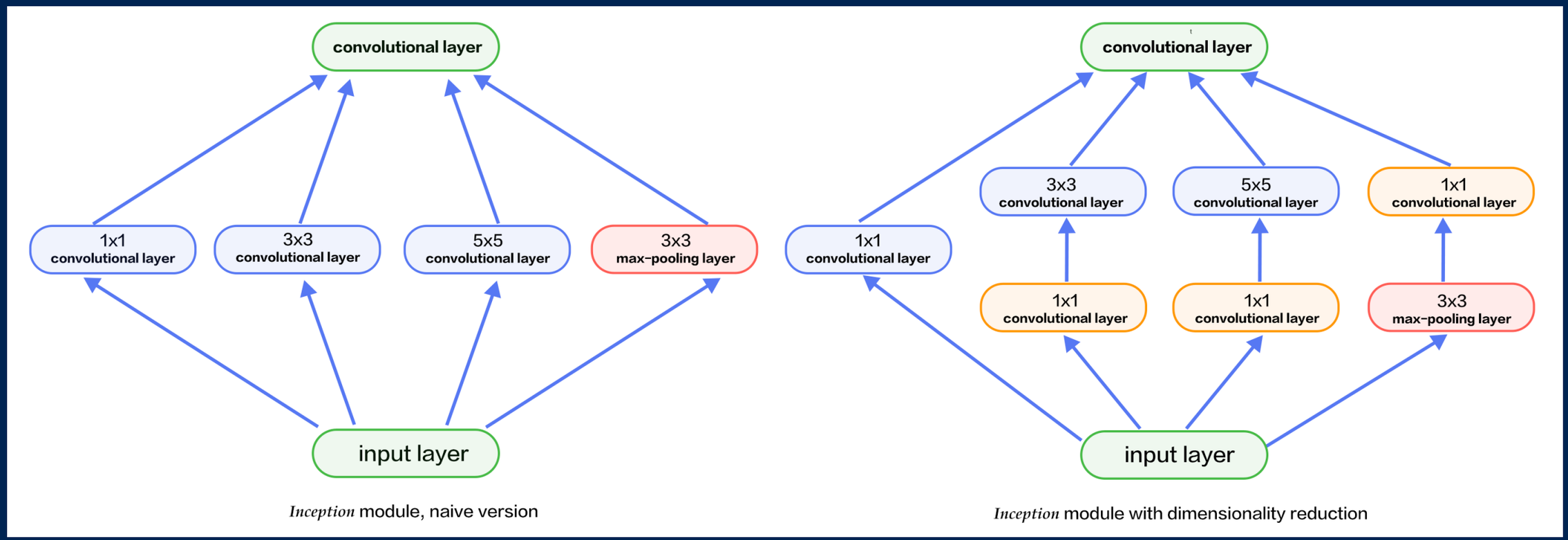
combined: vggnet





inception node

- **parallel execution, 1x1 convolution**

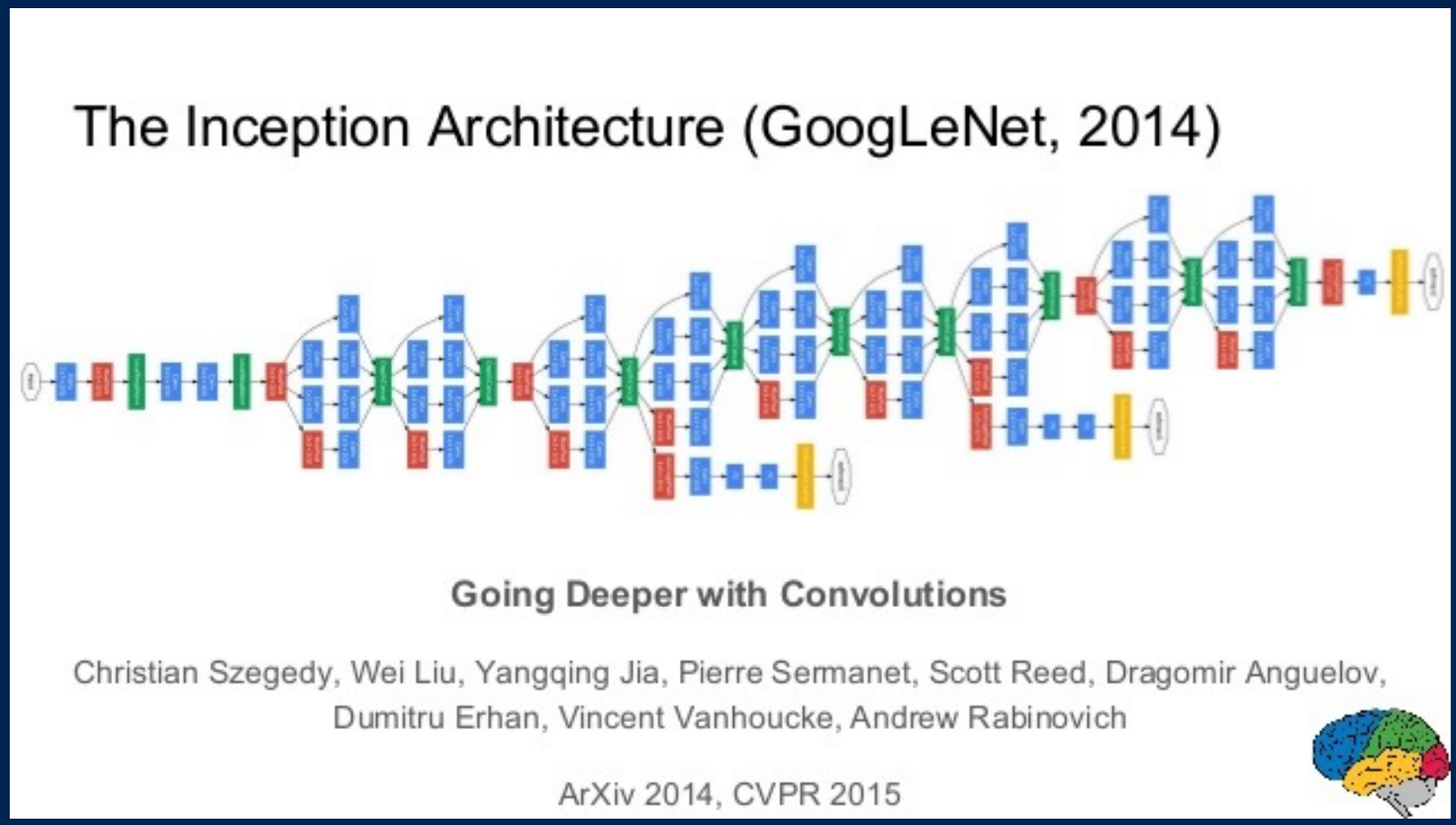


Inception module, naive version

Inception module with dimensionality reduction

- **iamaaditya.github.io/2016/03/one-by-one-convolution/**

inception graph



- hacktilldawn.com/2016/09/25/inception-modules-explained-and-implemented/

walk the path

training

- -1) python, roulette
- 0) calc 1, linear algebra basics, geometry
- 1) coursera, keras + tensorflow
- 2) fast.ai 2018 sequence, pytorch
- 3) read, practice, get into real world

platform

- 0) cpu + tools
- 1) basic vm's (coursera, paperspace)
- 2) cloud software (aws, gcp)
- 3) edge: mobile devices/embedded
- 4) custom hardware (tpu, volta, asic)

know your enemy

questions

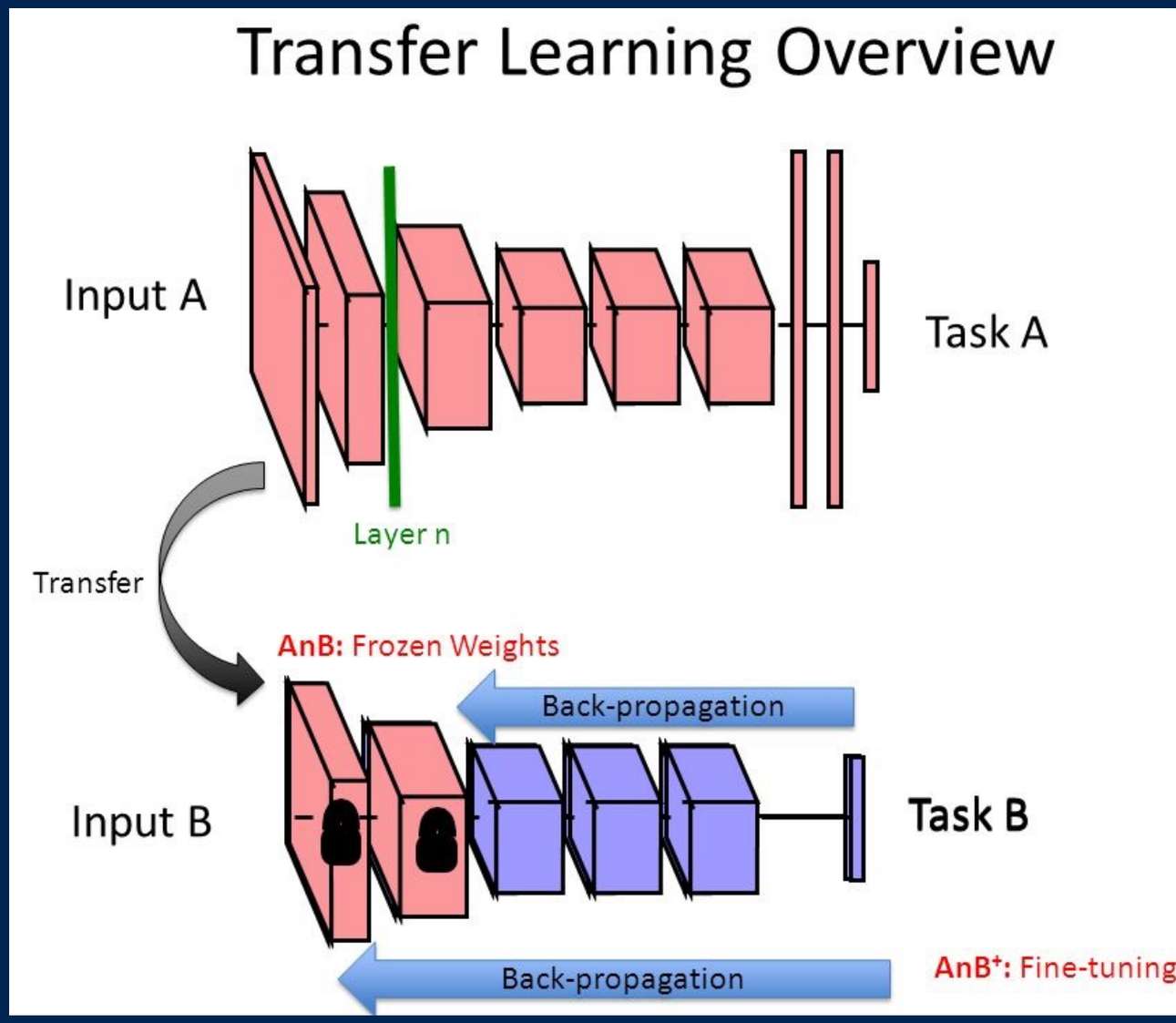
- what problem are we trying to solve?
- why are we trying to solve it?
- what data do we have?
- what does a right answer look like?

data

- more data makes most machine learning problems easier
- augmentation
- one-off solutions to get more data
- get creative and don’t do things right

model retraining

- **let's not rebuild our model from scratch!**
- **can reuse existing model**
- **re-run training on part of model with new data set**



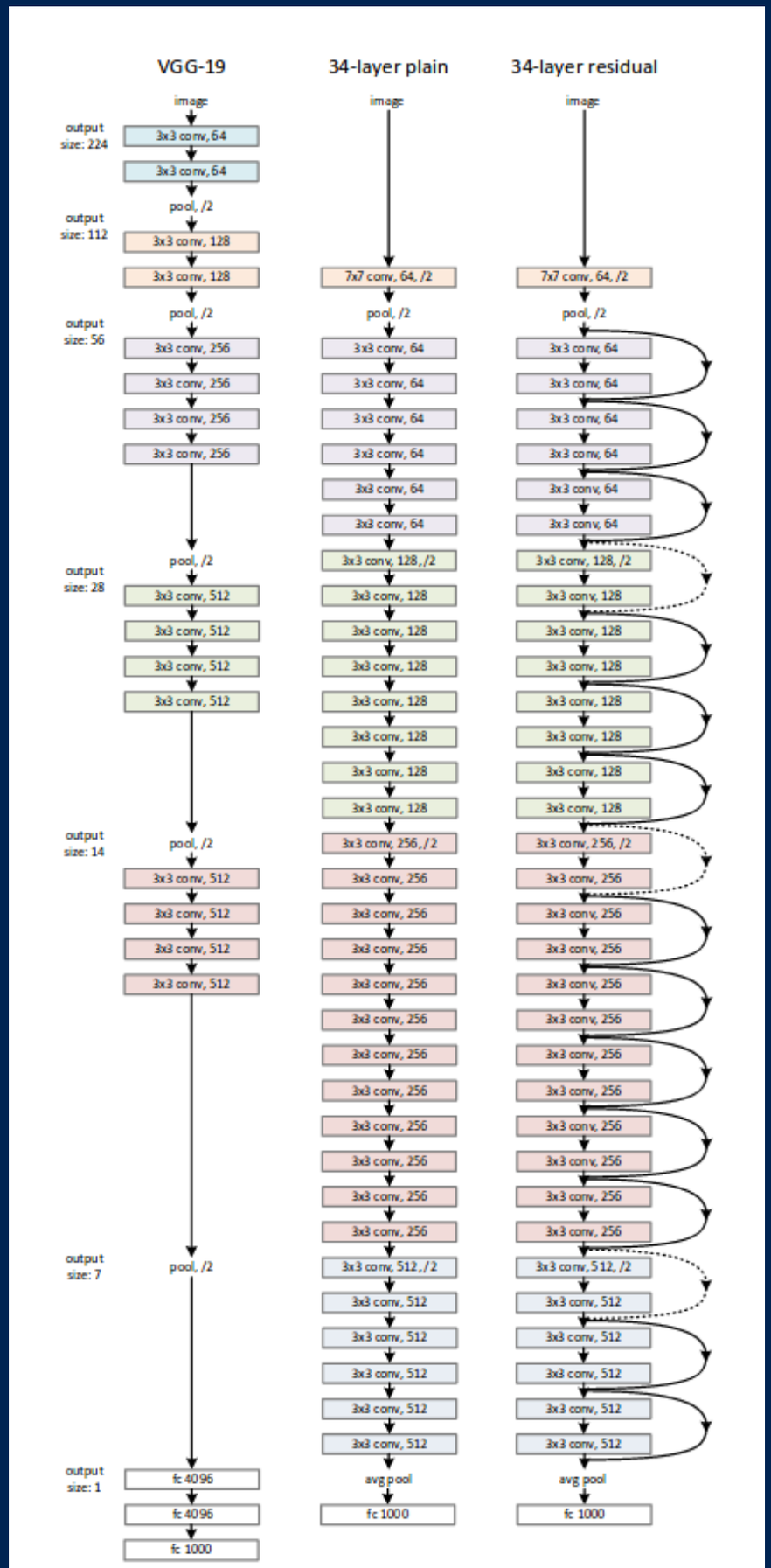
prepare your mind

image recognition

- 2017: basic nn, vgg, inception, resnet, mobilenets
- 2018: densenet, wide resnet, darknet, nasnet, amoebanet
- 2018 mobile: squeezenet, shufflenet, nasnet mobile, mobilenets v2

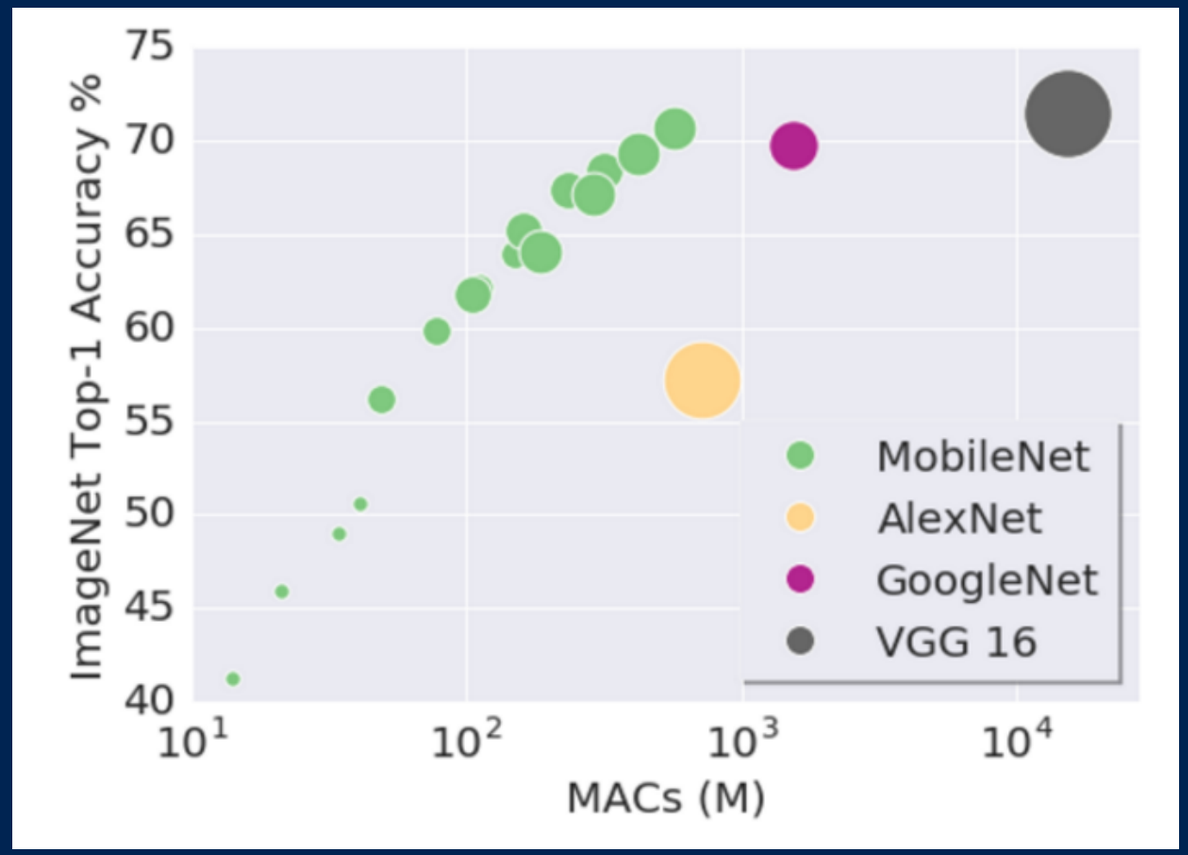
resnet

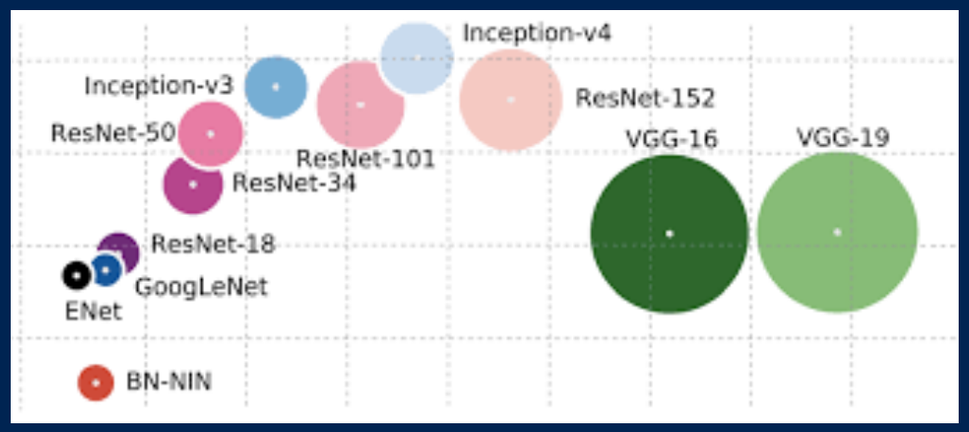
- residual networks
- skip connections
- even deeper training
- turicreate resnet demo



mobilenets

- depthwise separable convolutions
- announced april 17, paper, demo
- tf-lite demo
- february 18: v2





object detection

- yolo, yolo 9k, yolo v3
- vgg + ssd, resnet + ssd + retinanet
- mobilenets + ssd + tensorflow demo
- turicreate + yolo demo
- gcp + tensorflow for swift demo

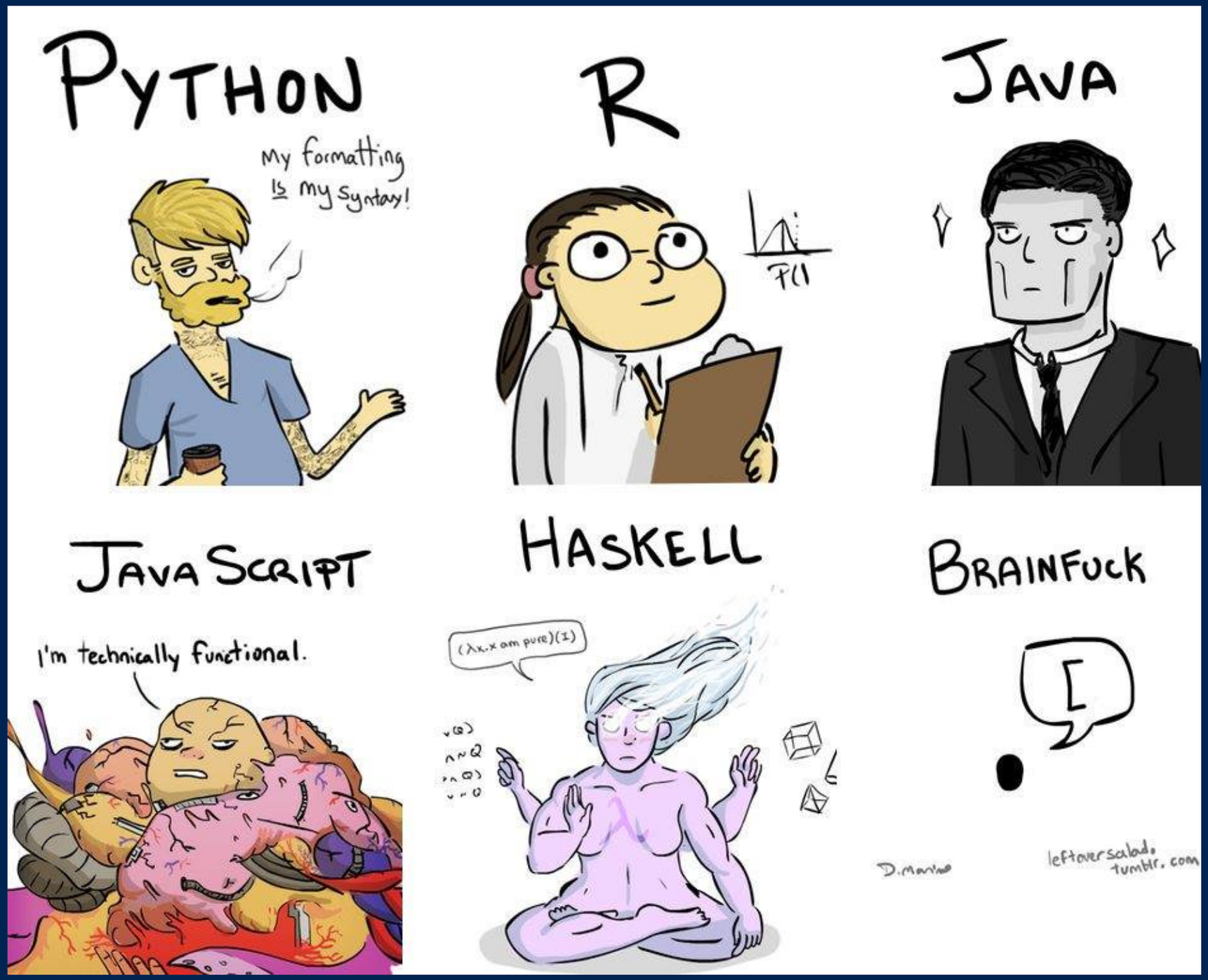
style transfer

- paper
- coursera example
- fast.ai example
- pytorch + onnx + coreml demo
- tensorflow android demo

super resolution

- paper
- japan demo
- pytorch onnx demo

contenders


PYTHON
My formatting is my syntax!
R
JAVA
JAVASCRIPT
I'm technically functional.
HASKELL
(λx.x am pure)(I)
BRAINFUCK
D.Marino
leftoversalad.tumblr.com

tensorflow + js

- tensorflow dev summit
- talk online shortly
- fun stuff
- keras bridge

tensorflow + swift

- potentially future of language/etc
- lattner + google —> cool stuff
- tpu demo

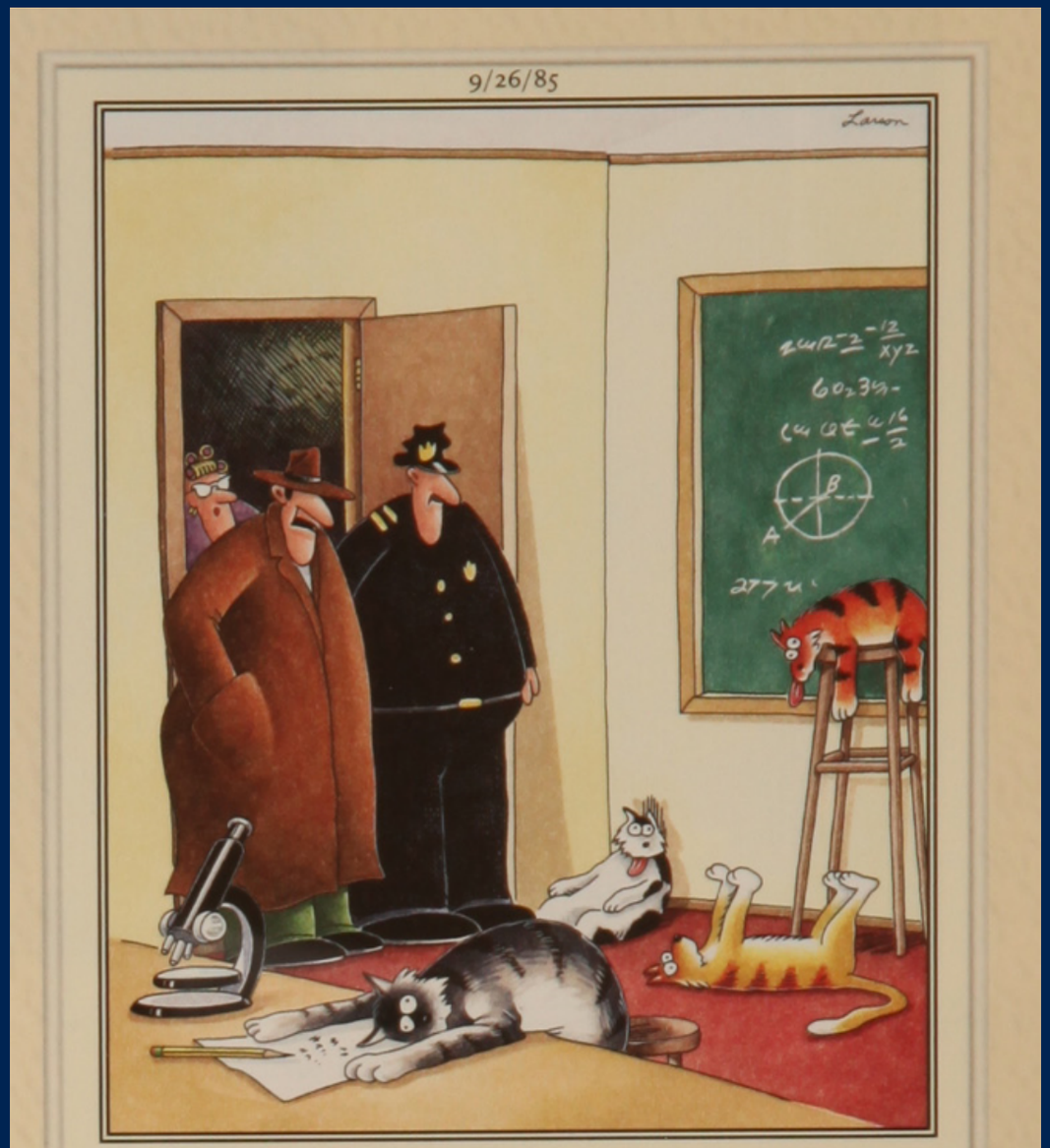
caffe

- potentially future of language/etc
- facebook backing
- coming soon
- api compatible with pytorch

firebase + mlkit

- **hybrid client/server approach**
- **build tf-lite model, upload to cloud, run on device**

look to the wind



"Notice all the computations, theoretical scribblings, and lab equipment, Norm. ... Yes, curiosity killed these cats."

create ml

- apple wrapping the whole enchilada
- drinking the ocean

tensor comprehensions

- multi gpu programming sucks
- need hard-typed language to make assumptions
- use complete graph model —> build hardware dsl for operations

off the beaten path

- audio/video segmentation
- language models
- gans
- game theory
- ...