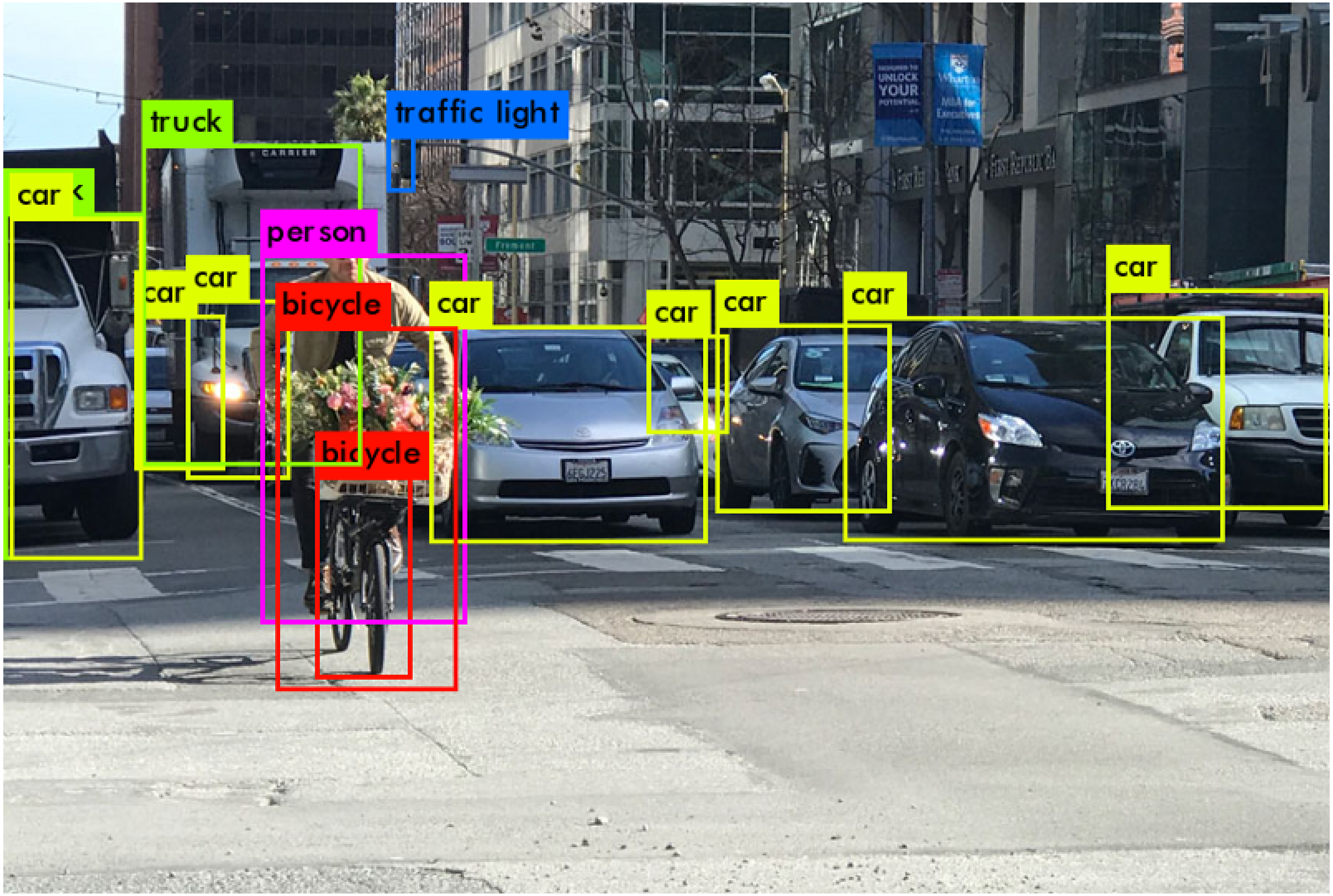
Deep Computer Vision



Who am I?

- PhD student – astronomy, instrumentation.
- MOOC junkie.
- Deep into deep learning.
- Offering training, workshops on DL.

Vikram Mark Radhakrishnan
radhakrishnan@strw.leidenuniv.nl

https://github.com/VikramRadhakrishnan/DeepCV

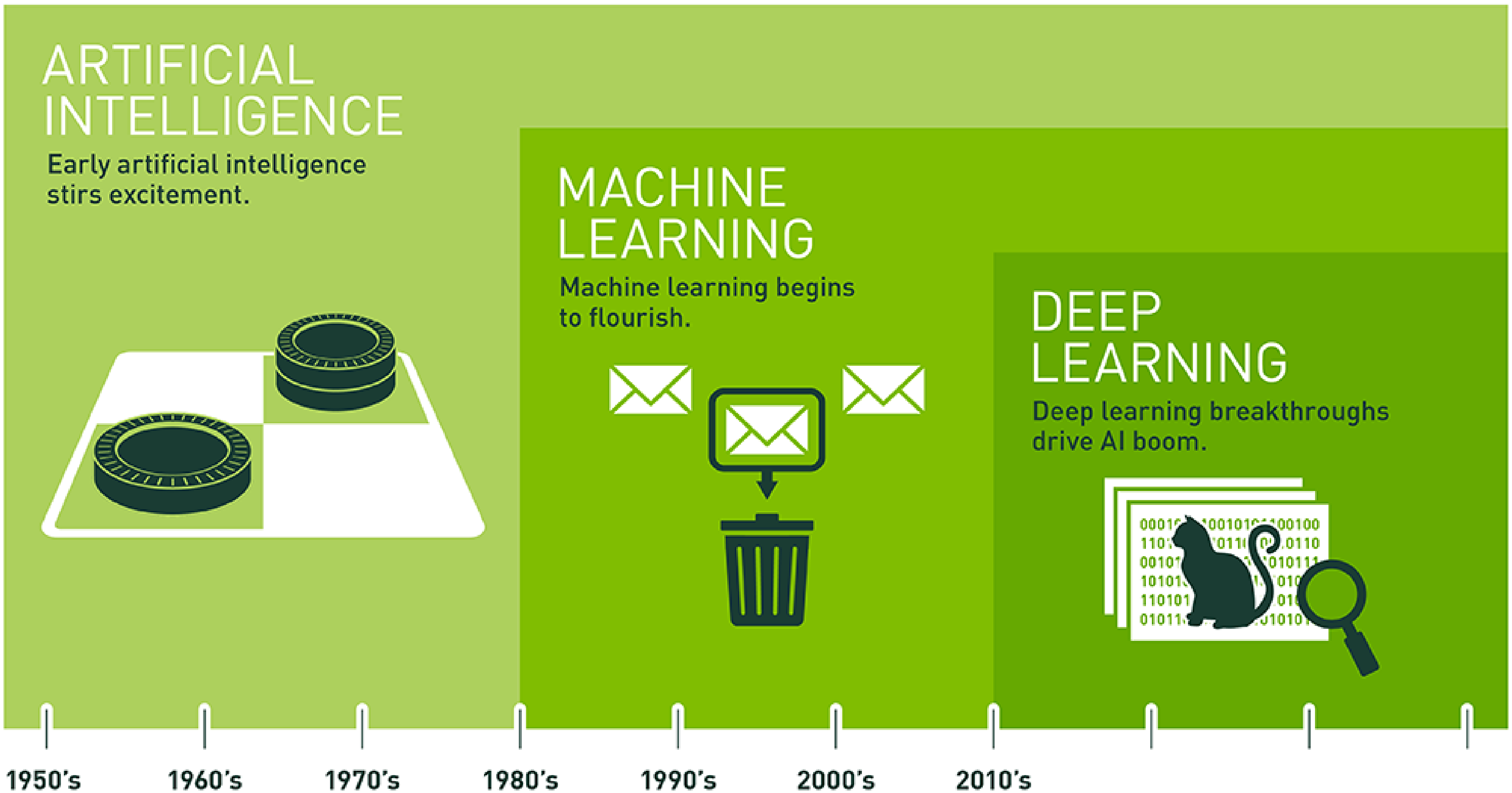
https://www.kaggle.com/c/digit-recognizer/kernels

Make computers see what we see

“Get” the picture

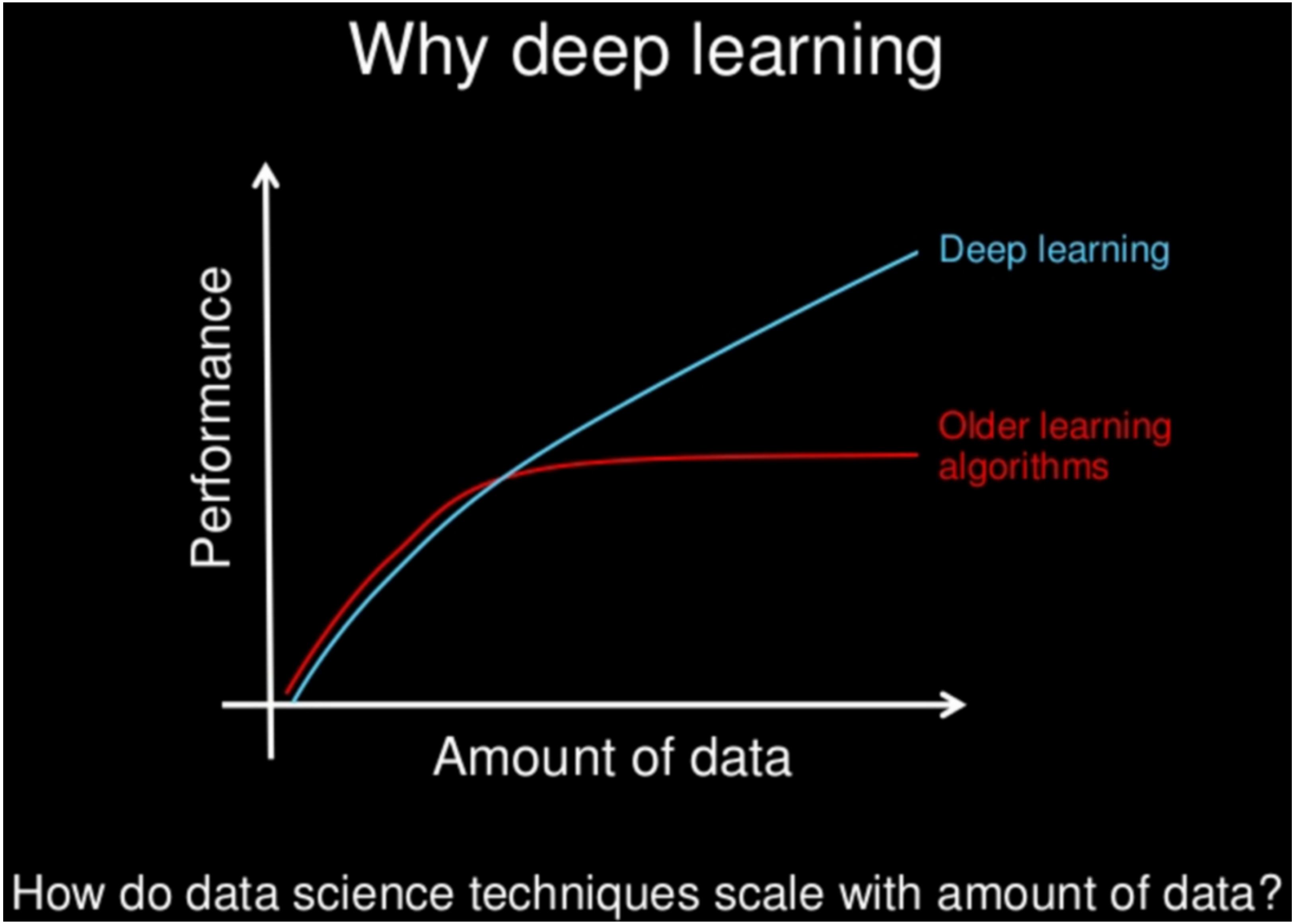
SO MUCH OF "AI" IS JUST FIGURING OUT WAYS TO OFFLOAD WORK ONTO RANDOM STRANGERS.

A.I. can be our visual cortex

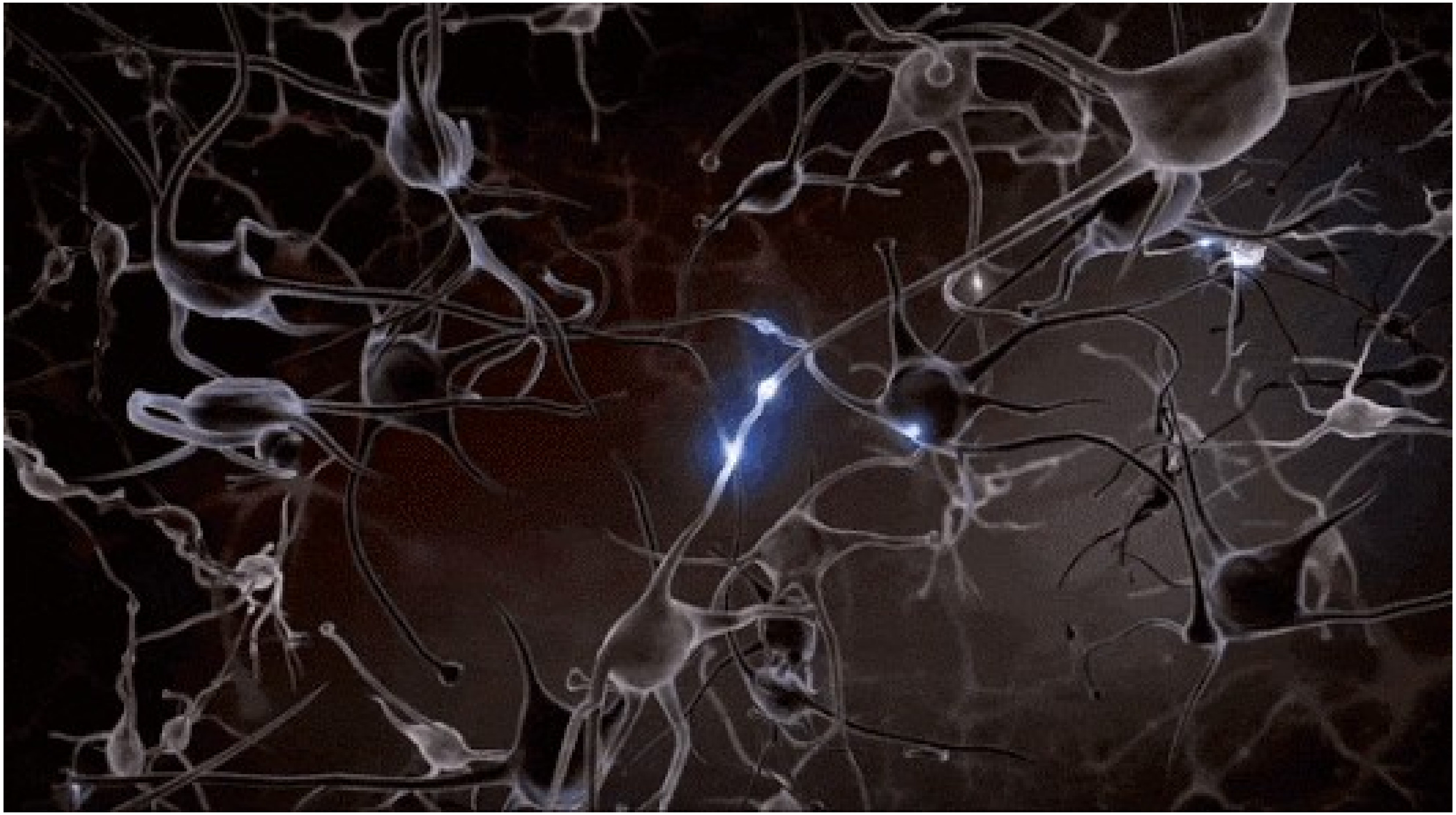


Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

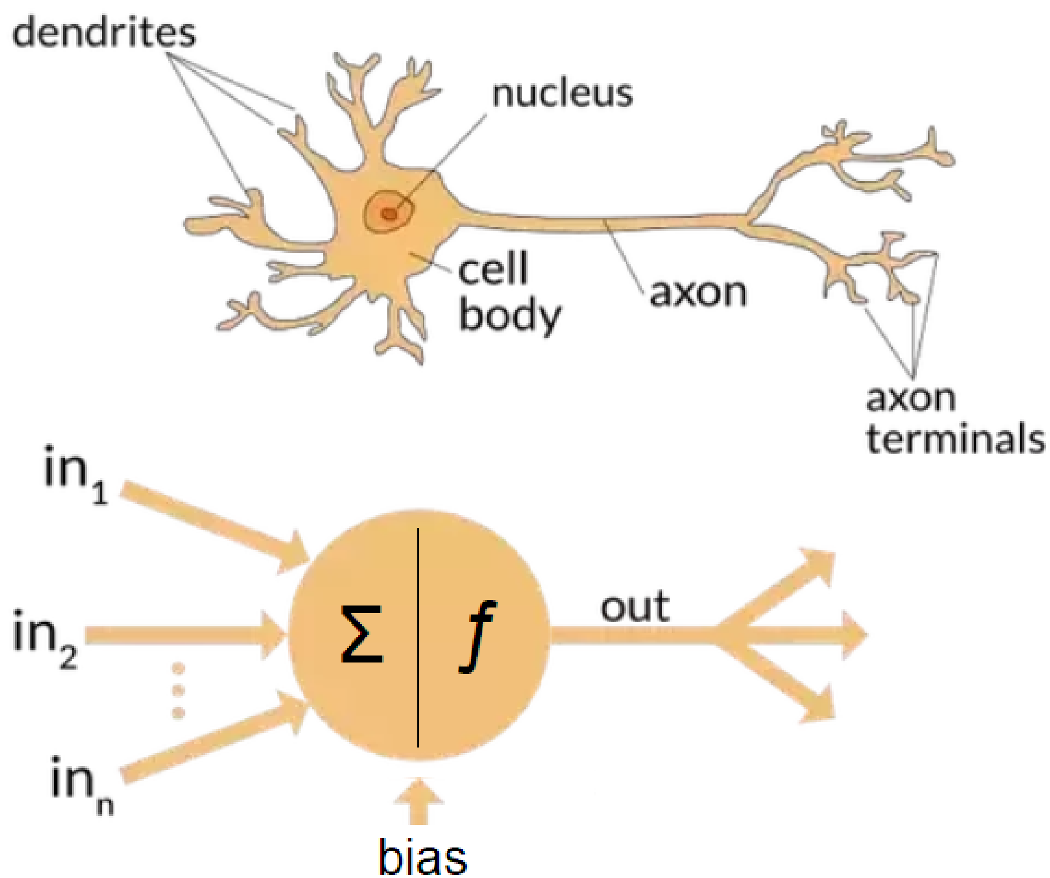
The dawn of deep learning



Based on biology

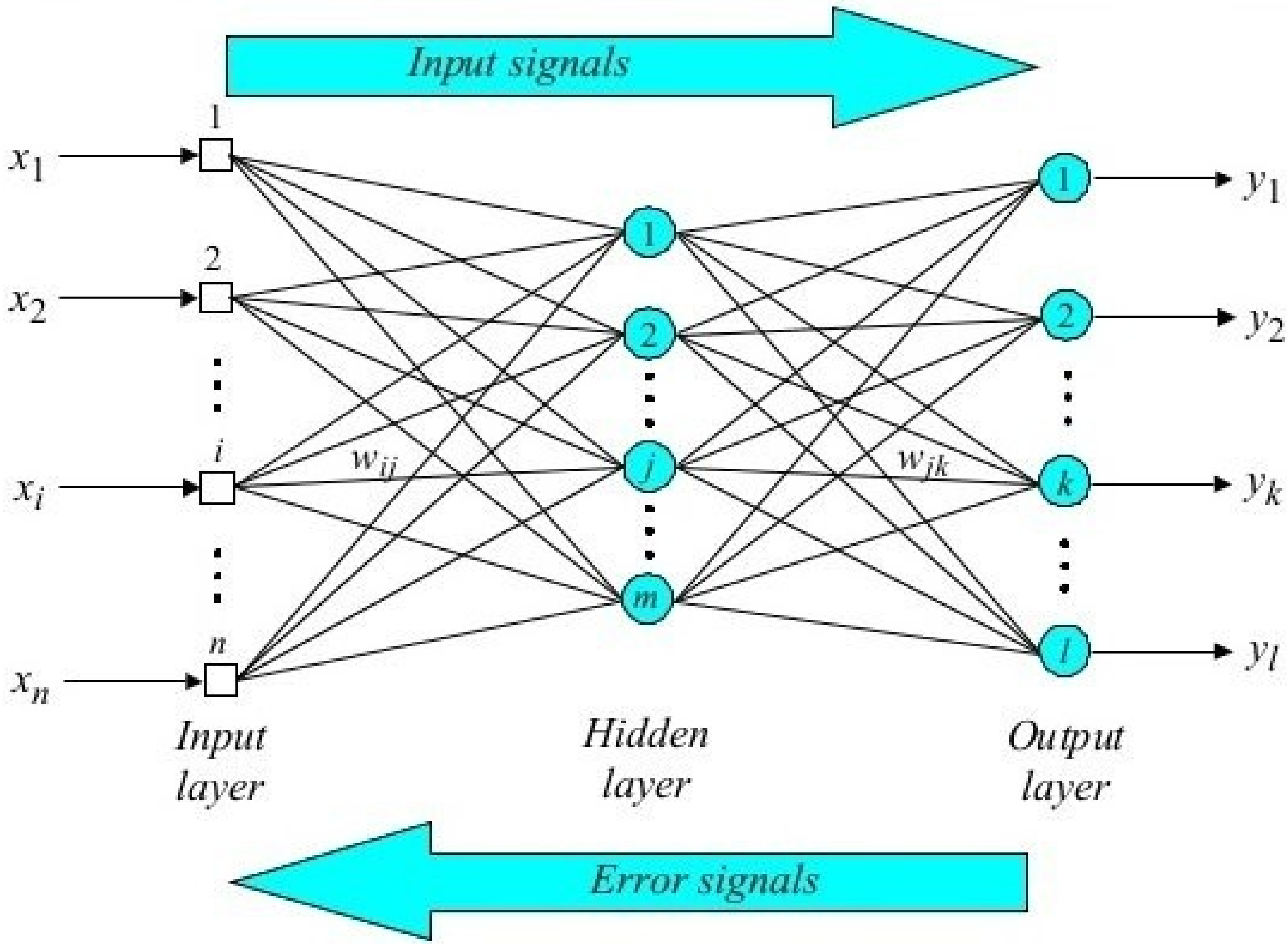


Neural networks - Home grown and in the wild



towardsdatascience.com

Fully connected neurons

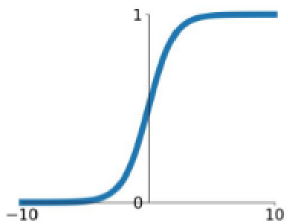


Activation functions

Activation Functions

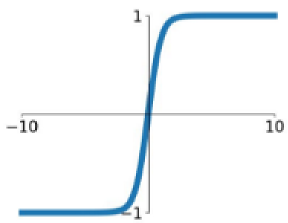
Sigmoid

$\sigma(x) = \frac{1}{1+e^{-x}}$



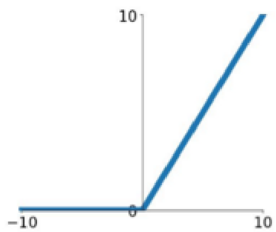
tanh

$\tanh(x)$



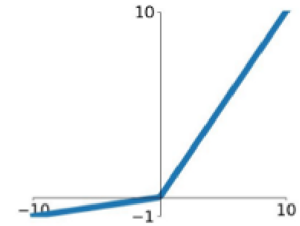
ReLU

$\max(0, x)$



Leaky ReLU

$\max(0.1x, x)$

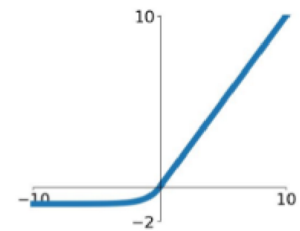


Maxout

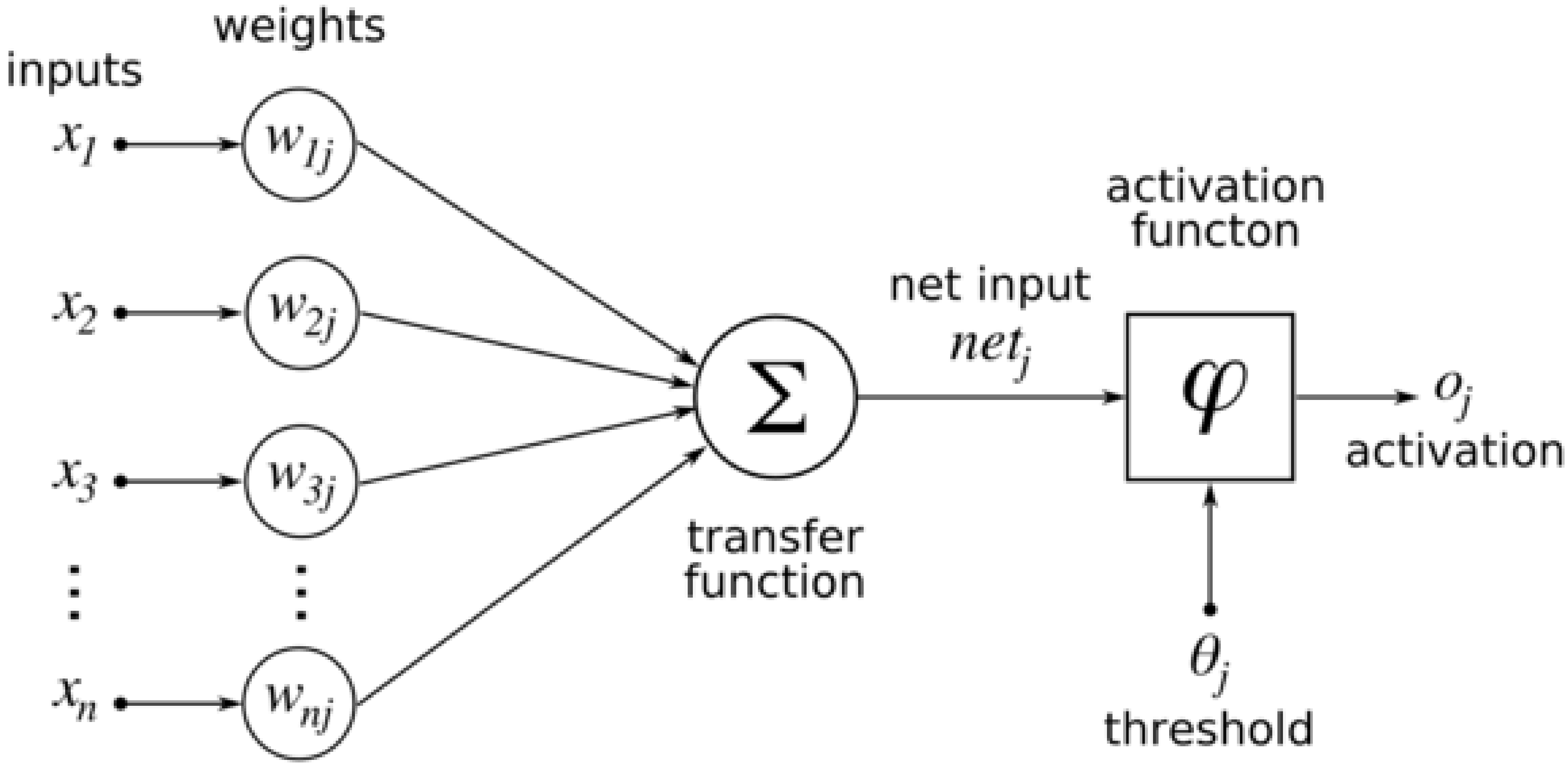
$\max(w_1^T x + b_1, w_2^T x + b_2)$

ELU

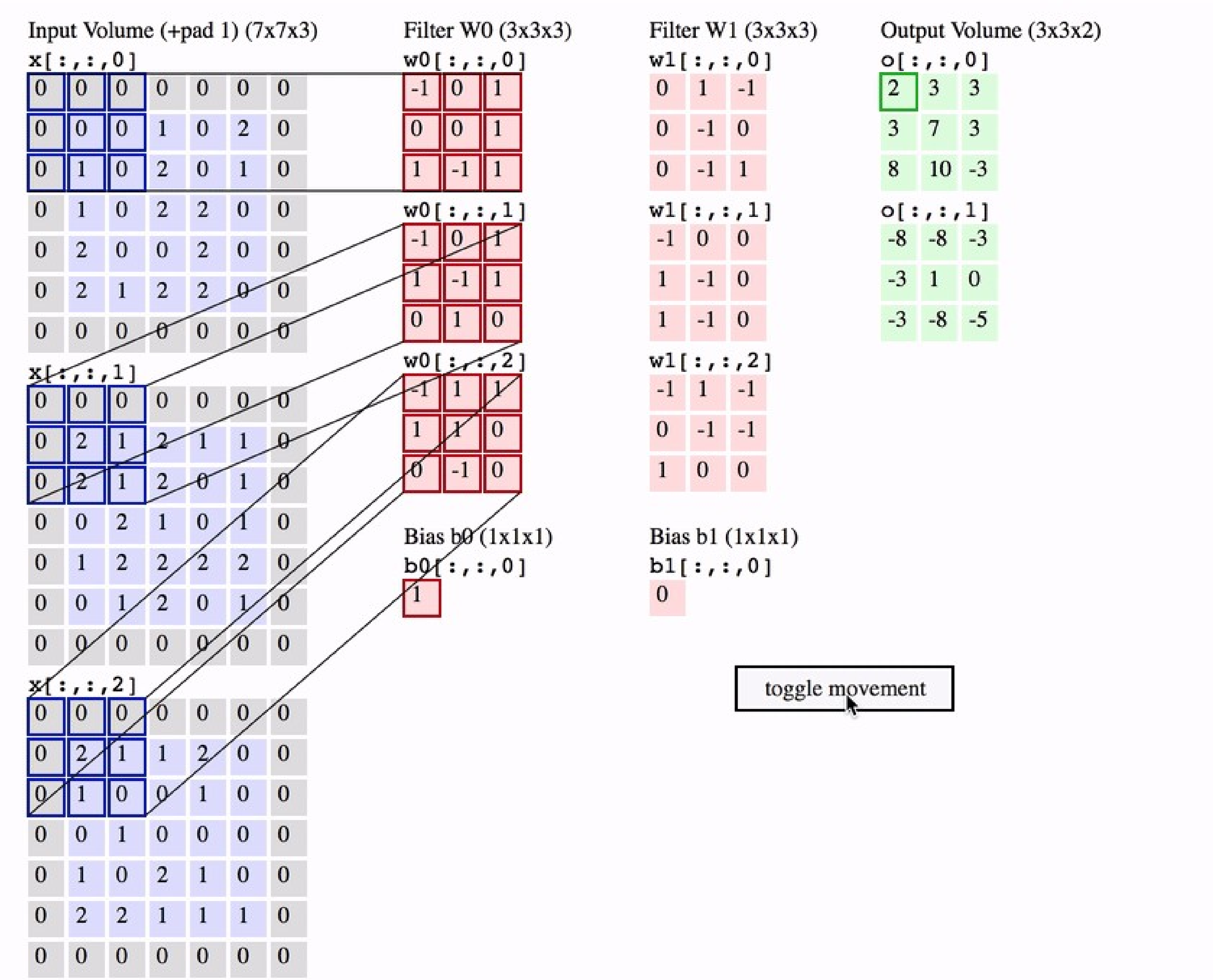
$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$



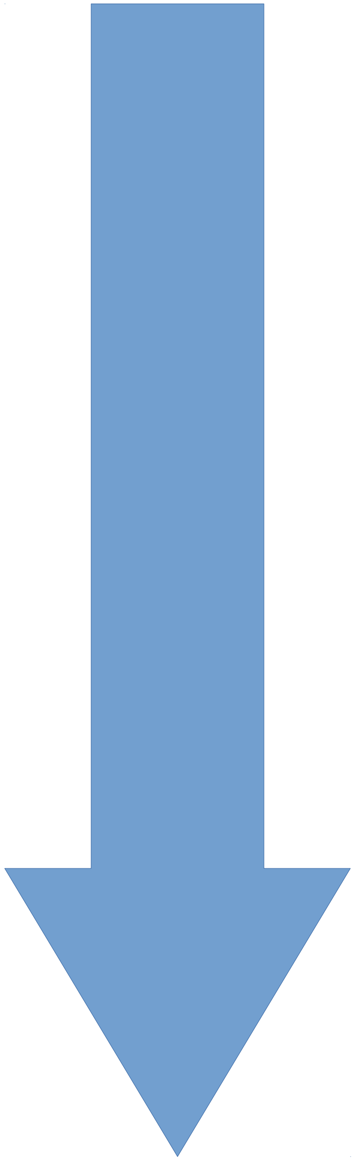
Linear → Non-linear

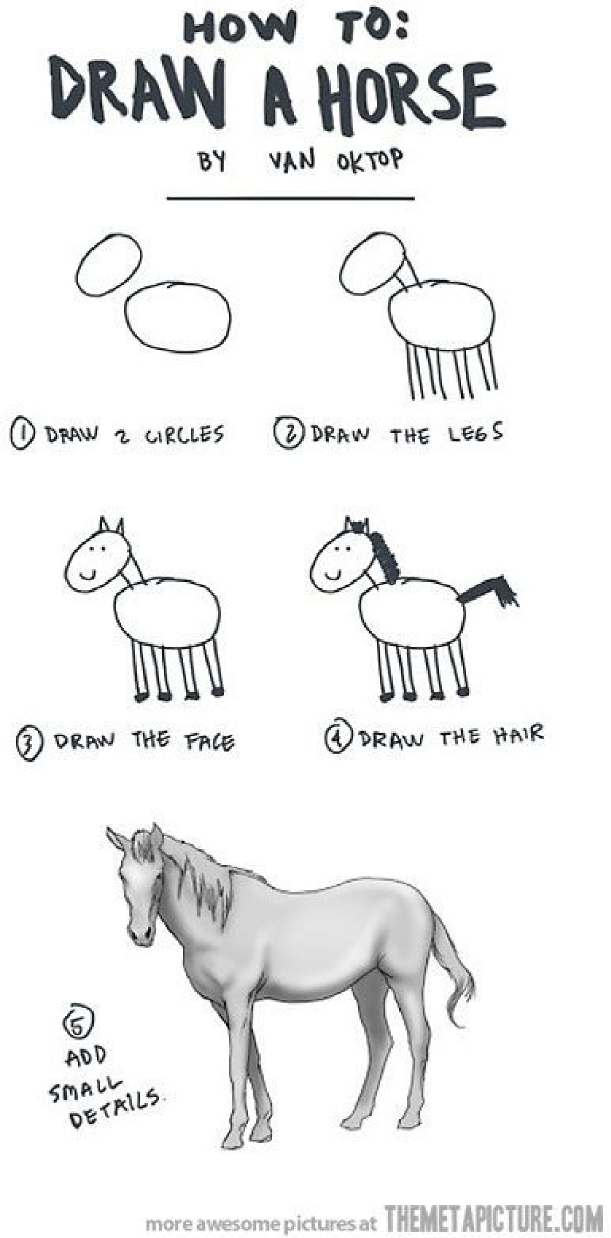


Convolution

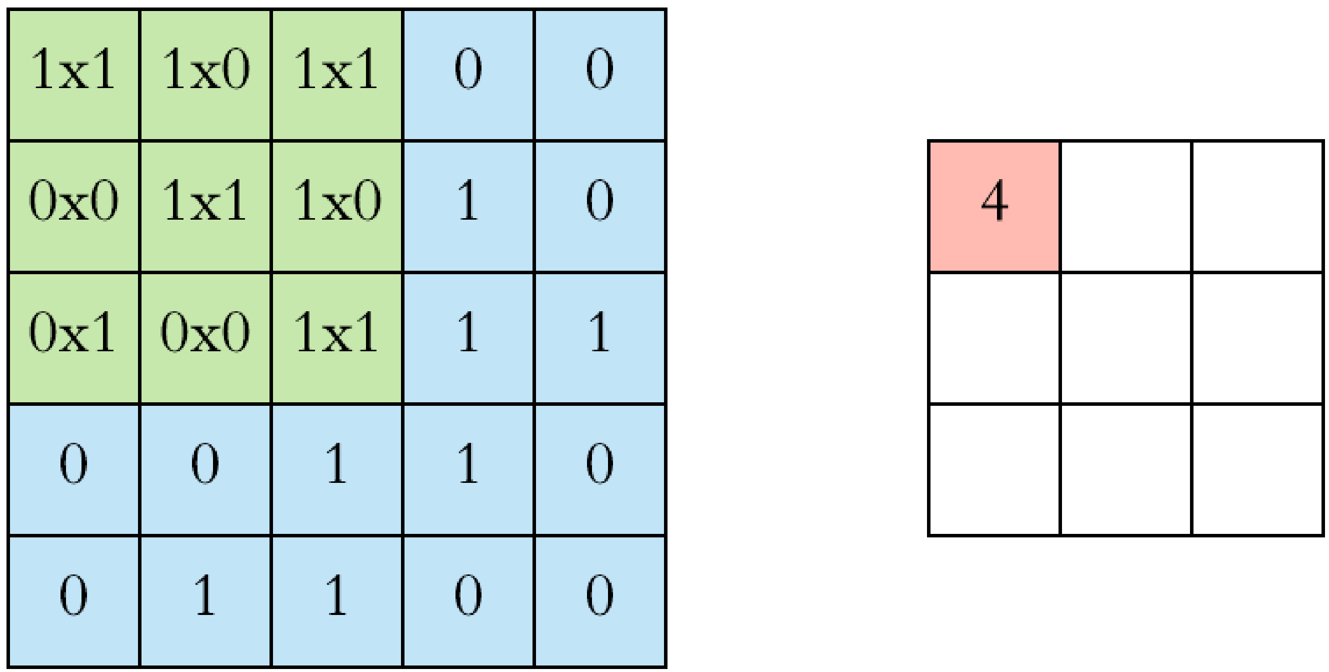


How our brains process images

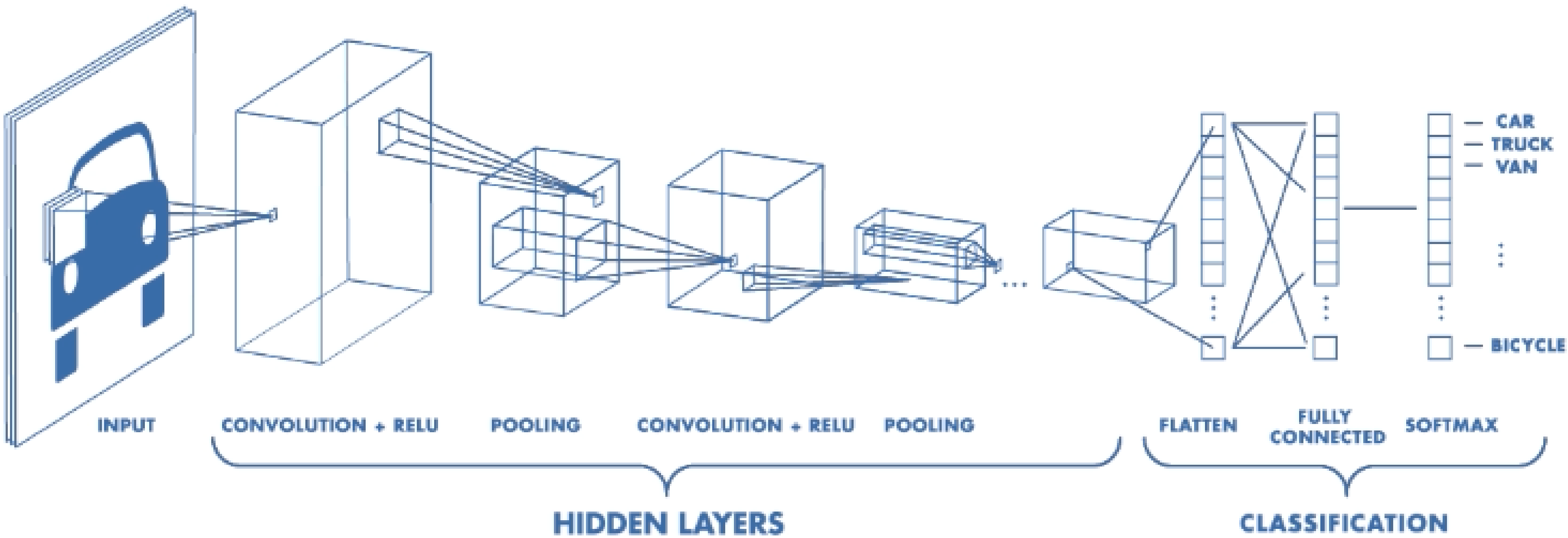




Pooling

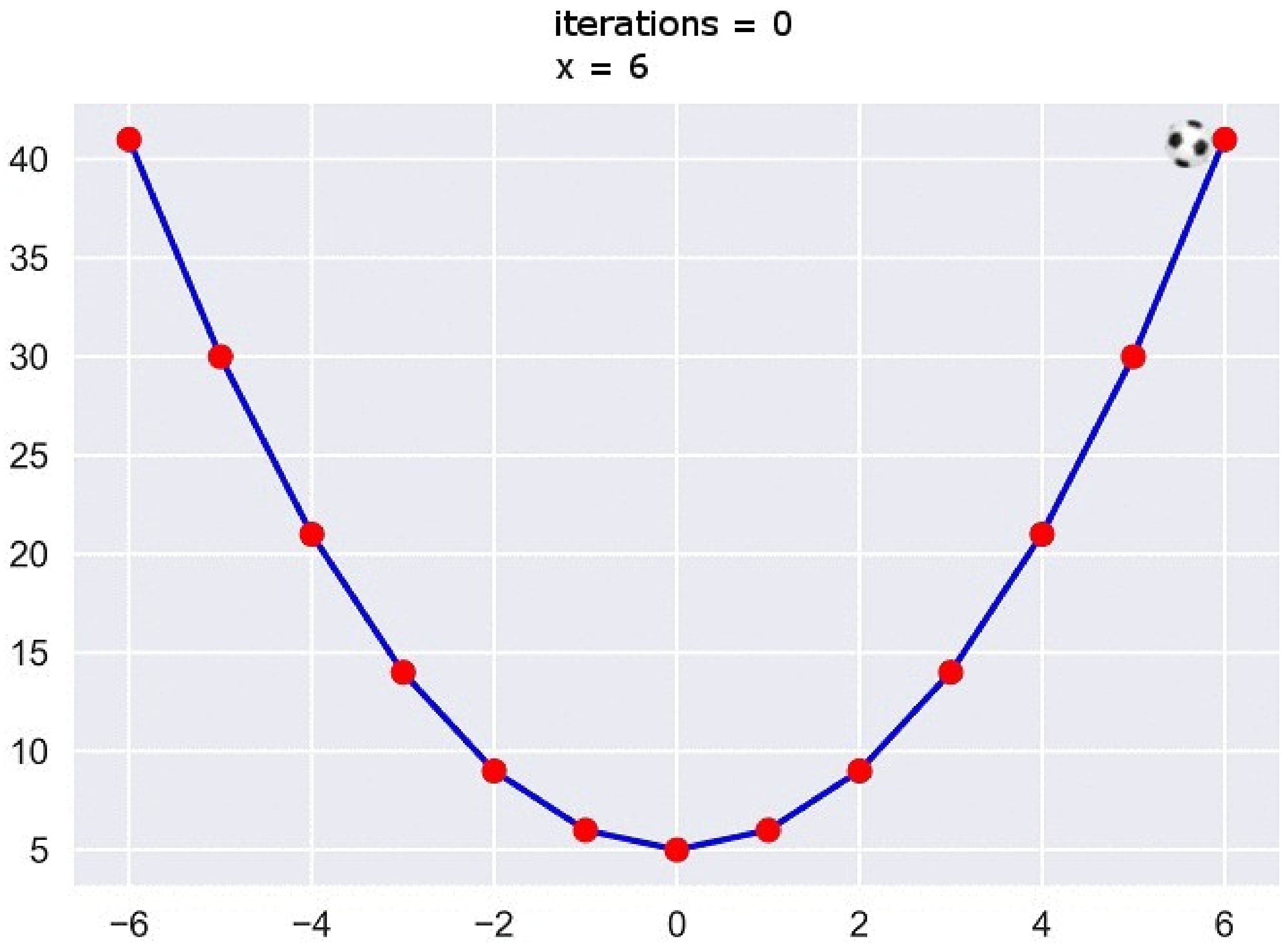


Putting them together

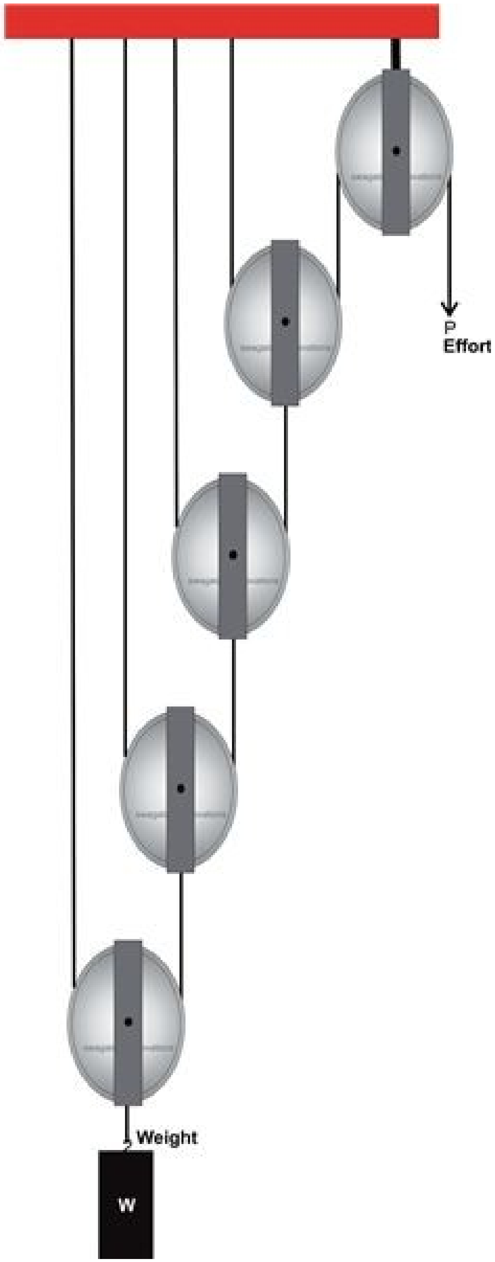


Learning – keep doing a better job

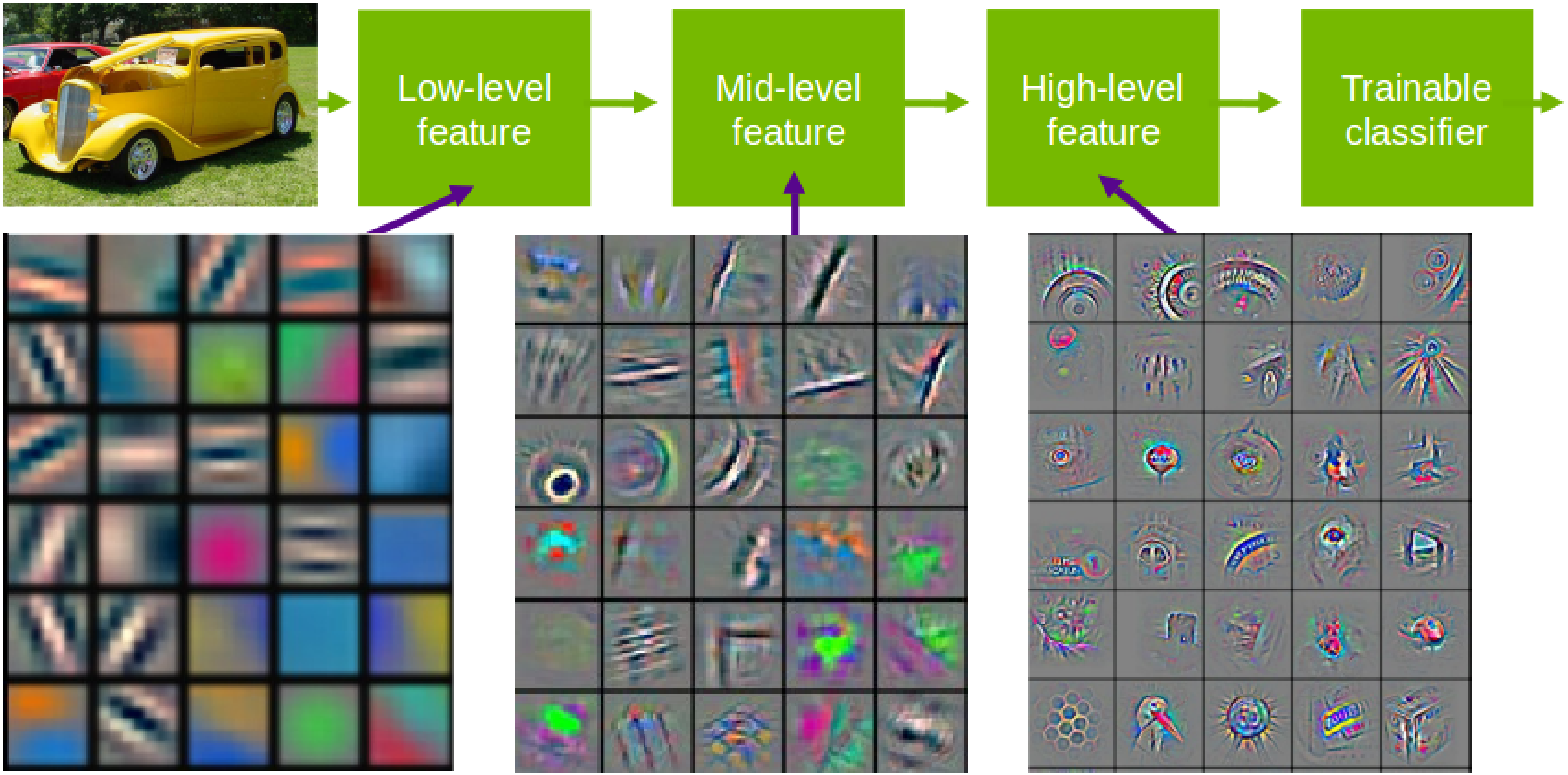
Gradient descent



Backprop!

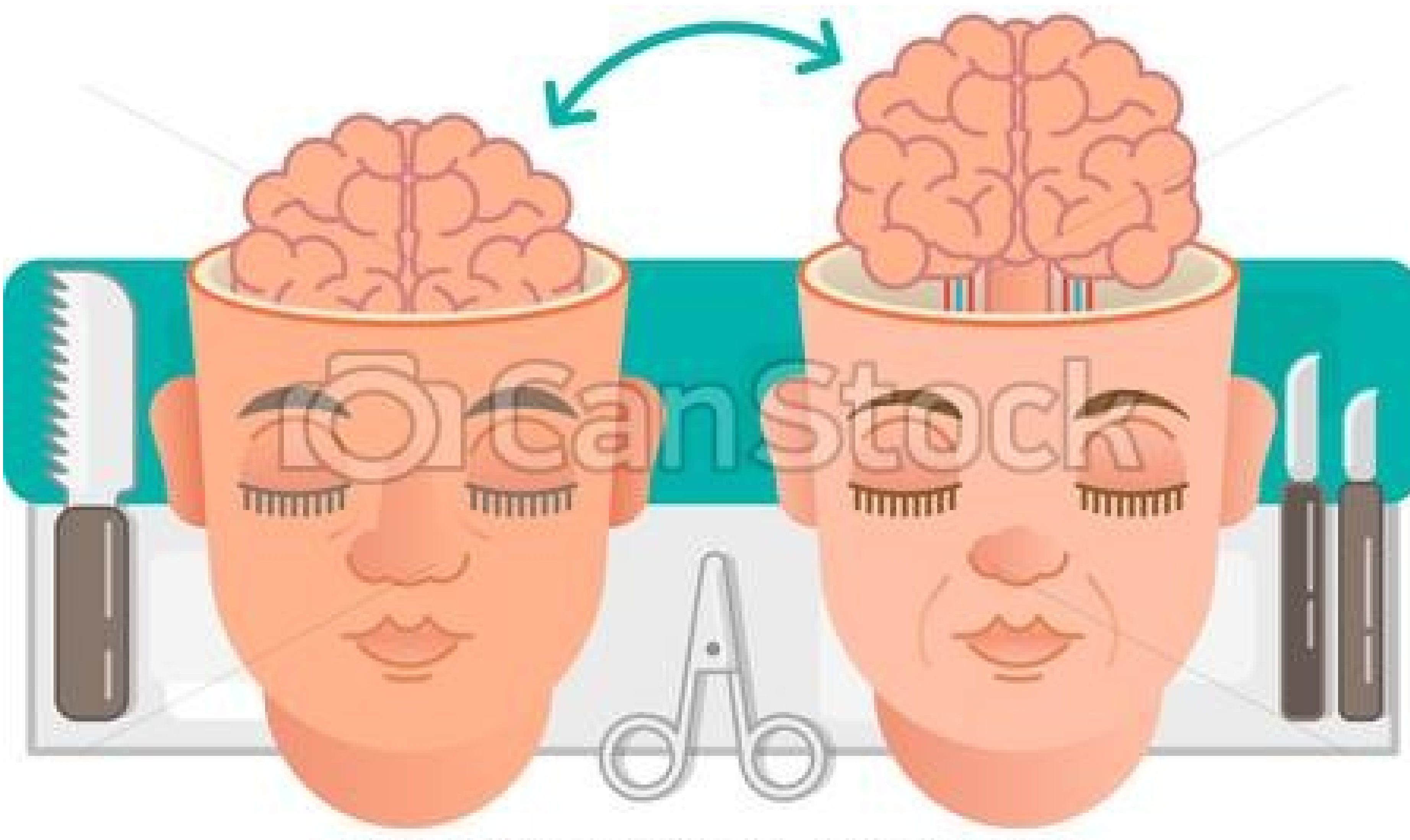


Feature extraction with CNN



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

Transfer learning



What can you do with this?

Convolutional neural networks

Transfer learning

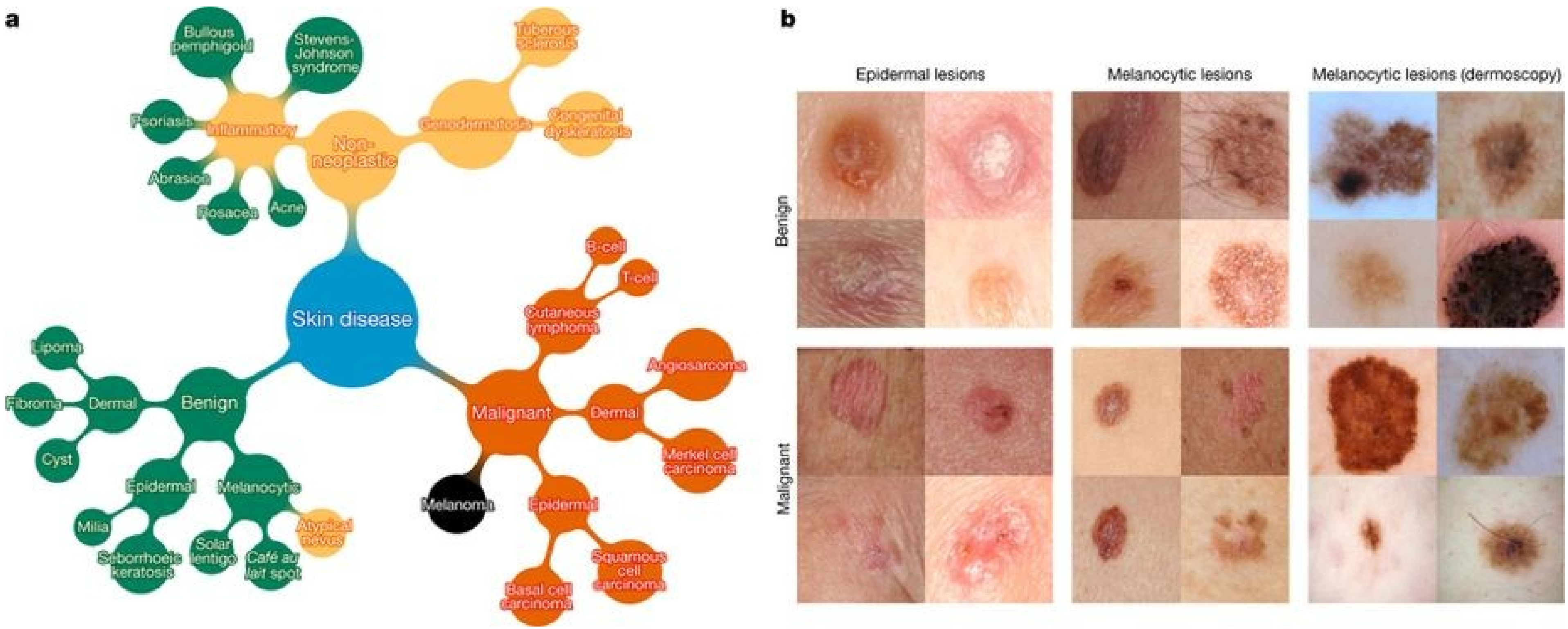
Advanced gradient descent algorithms

User friendly coding libraries (Keras, Fastai, Caffe2)

Access to powerful GPU hardware

Detect cancer better than a doctor

Accuracy of deep learning model > accuracy of 21 board-certified dermatologists

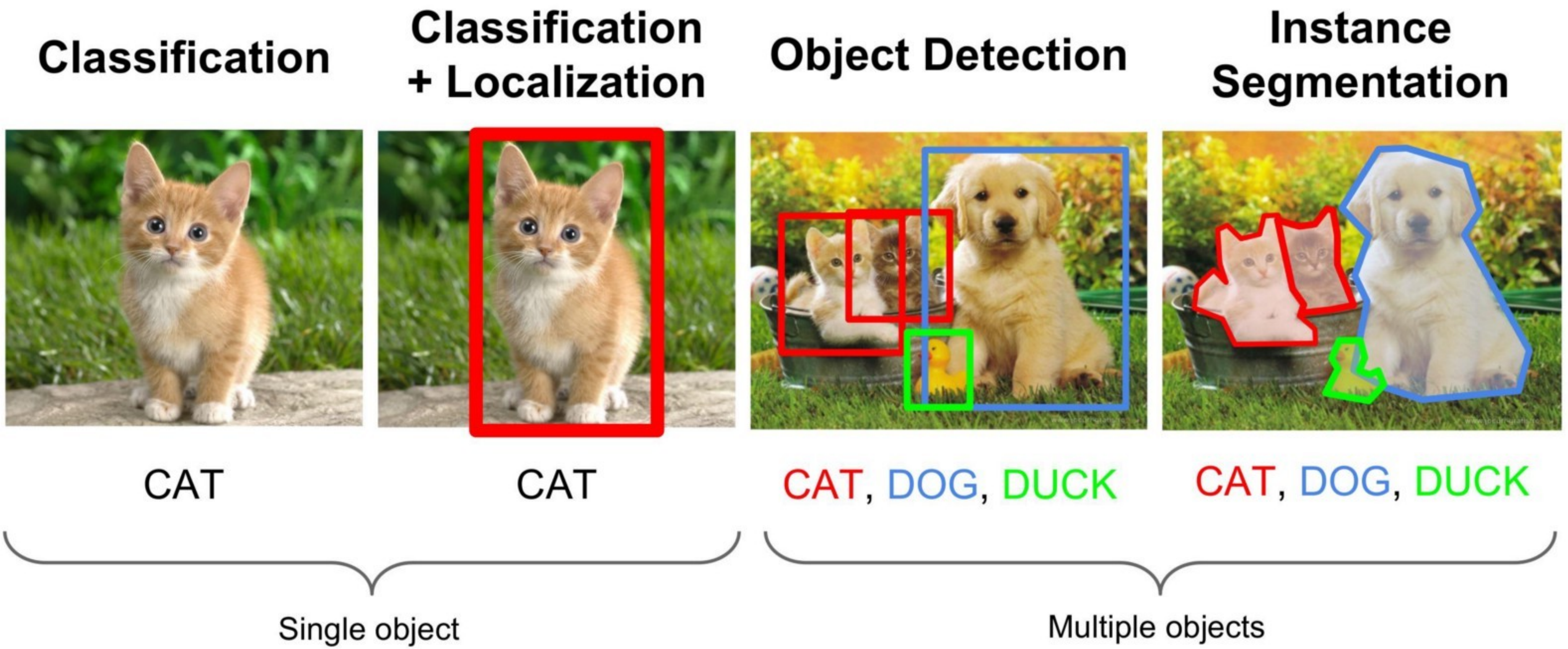


Esteva et al(2017) - Dermatologist-level classification of skin cancer with deep neural networks

https://www.nature.com/articles/nature21056

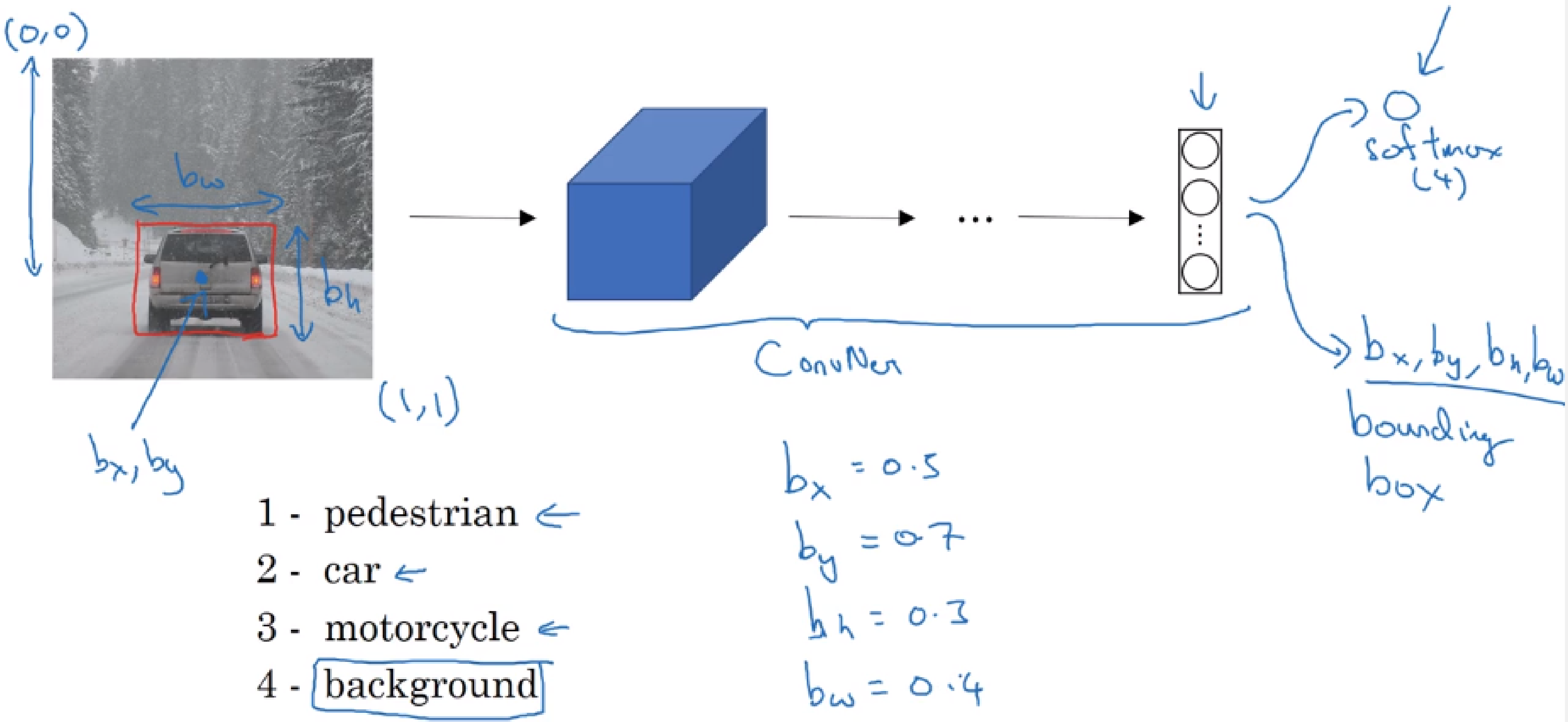
Create ridiculously expensive art

$432,000



Arthur Ouaknine -
https://medium.com/comet-app/

Bounding boxes



Credits to Andrew Ng

Bounding box examples

$$Y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

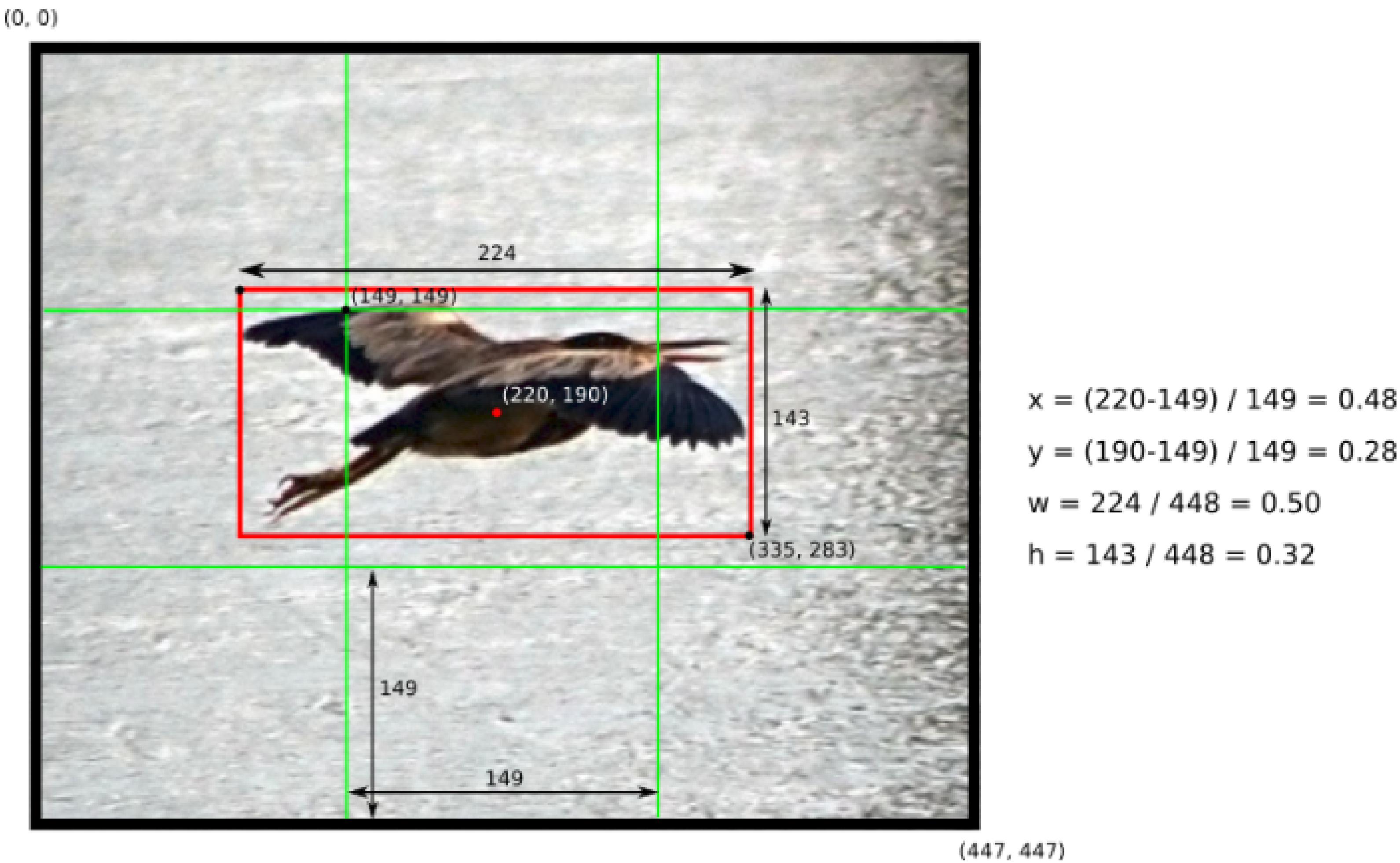
Detection

$$\begin{bmatrix} 1 \\ b_x \\ b_y \\ b_h \\ b_w \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

No detection

$$\begin{bmatrix} 0 \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \end{bmatrix}$$

Calculating box coordinates



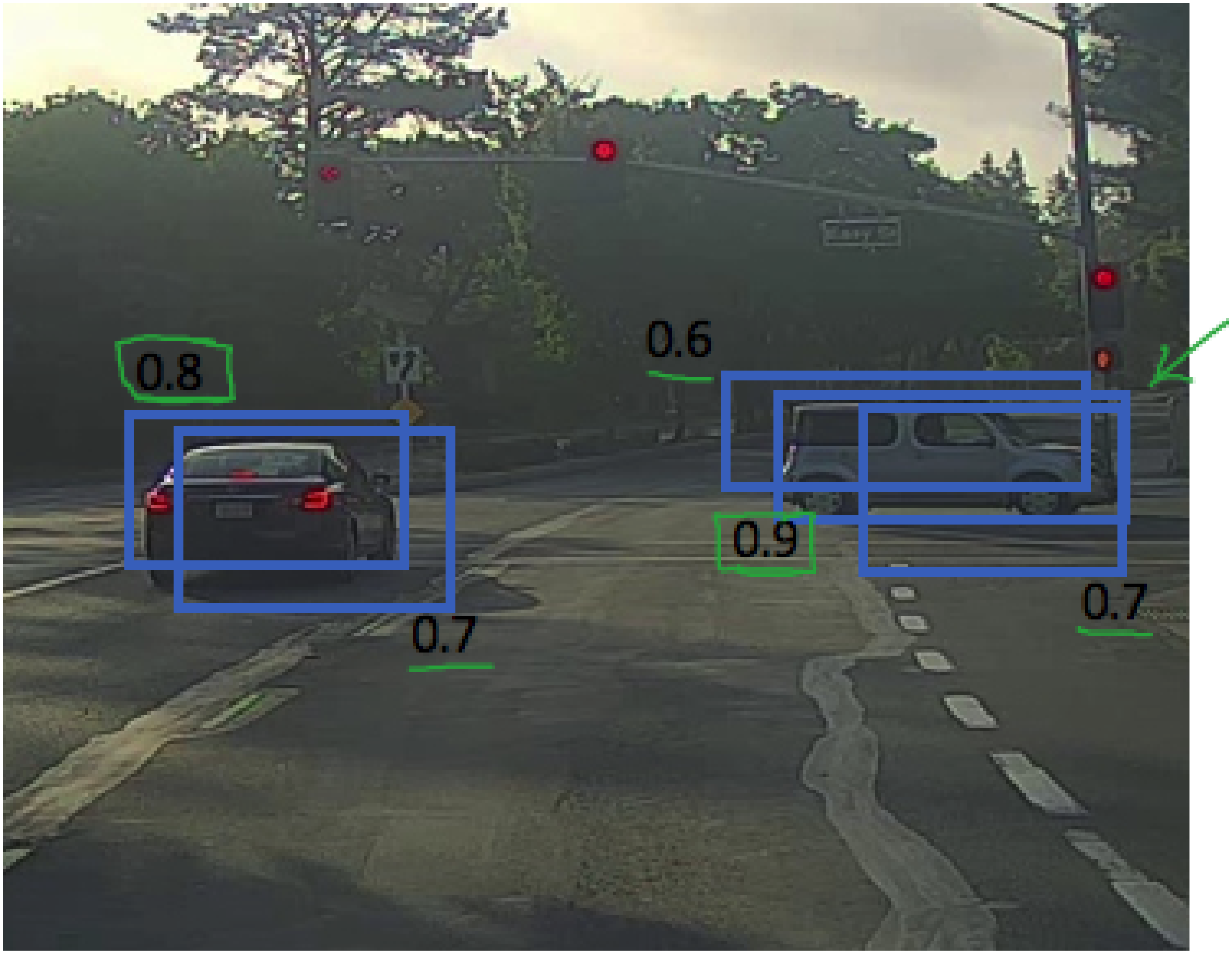
Example of how to calculate box coordinates in a 448x448 image with S=3. Note how the (x,y) coordinates are calculated relative to the center grid cell

https://hackernoon.com

IOU

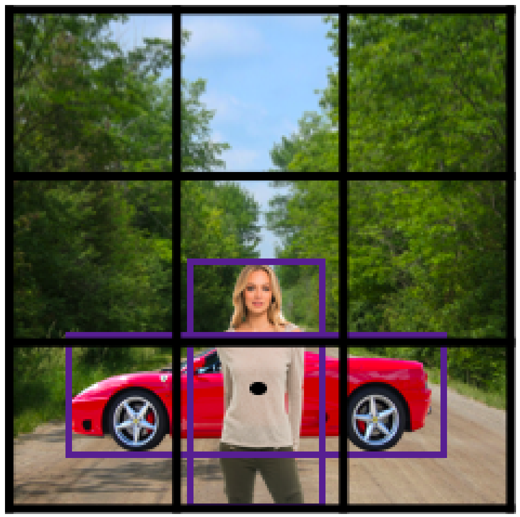
$$IoU = \frac{size\ of\ the\ intersection\ area}{size\ of\ the\ union\ area}$$

Non max suppression



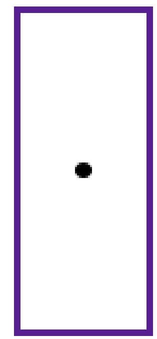
Anchor boxes

$$Y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \\ p'_c \\ b'_x \\ b'_y \\ b'_h \\ b'_w \\ c'_1 \\ c'_2 \\ c'_3 \end{bmatrix}$$

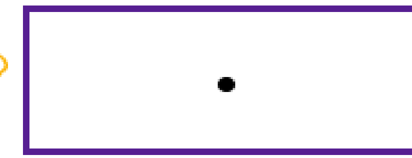


$$\mathbf{y} = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

Anchor box 1:



Anchor box 2:



$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \\ p_c \\ b_x \\ \vdots \\ c_3 \end{bmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \\ \end{array}\right\} \text{anchor box 1} \\ \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} \text{anchor box 2} \end{array}$$