MealoMi

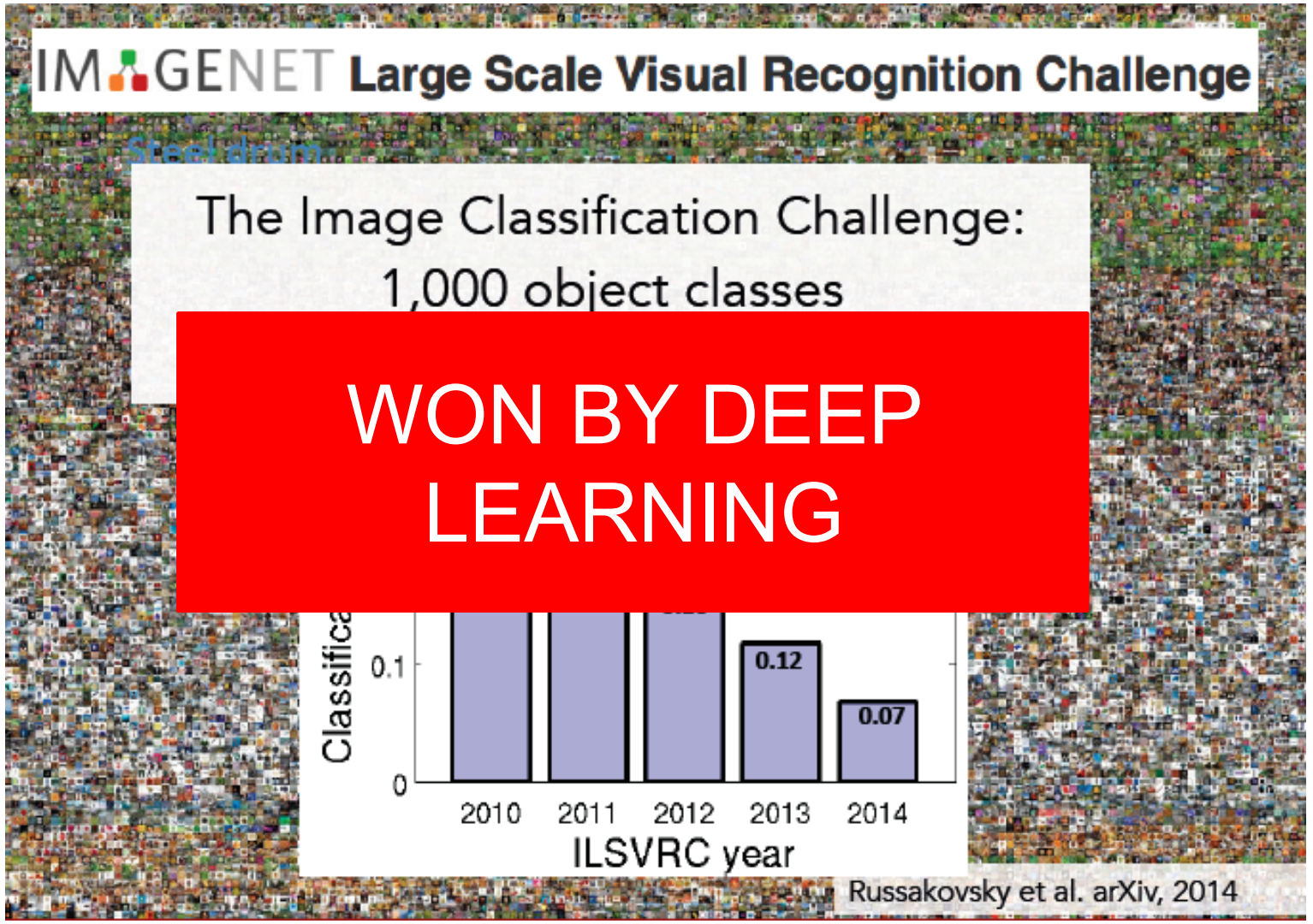
Introduction to Deep Learning

Patrick Christ

The development of full artificial intelligence could spell the end of the human race.

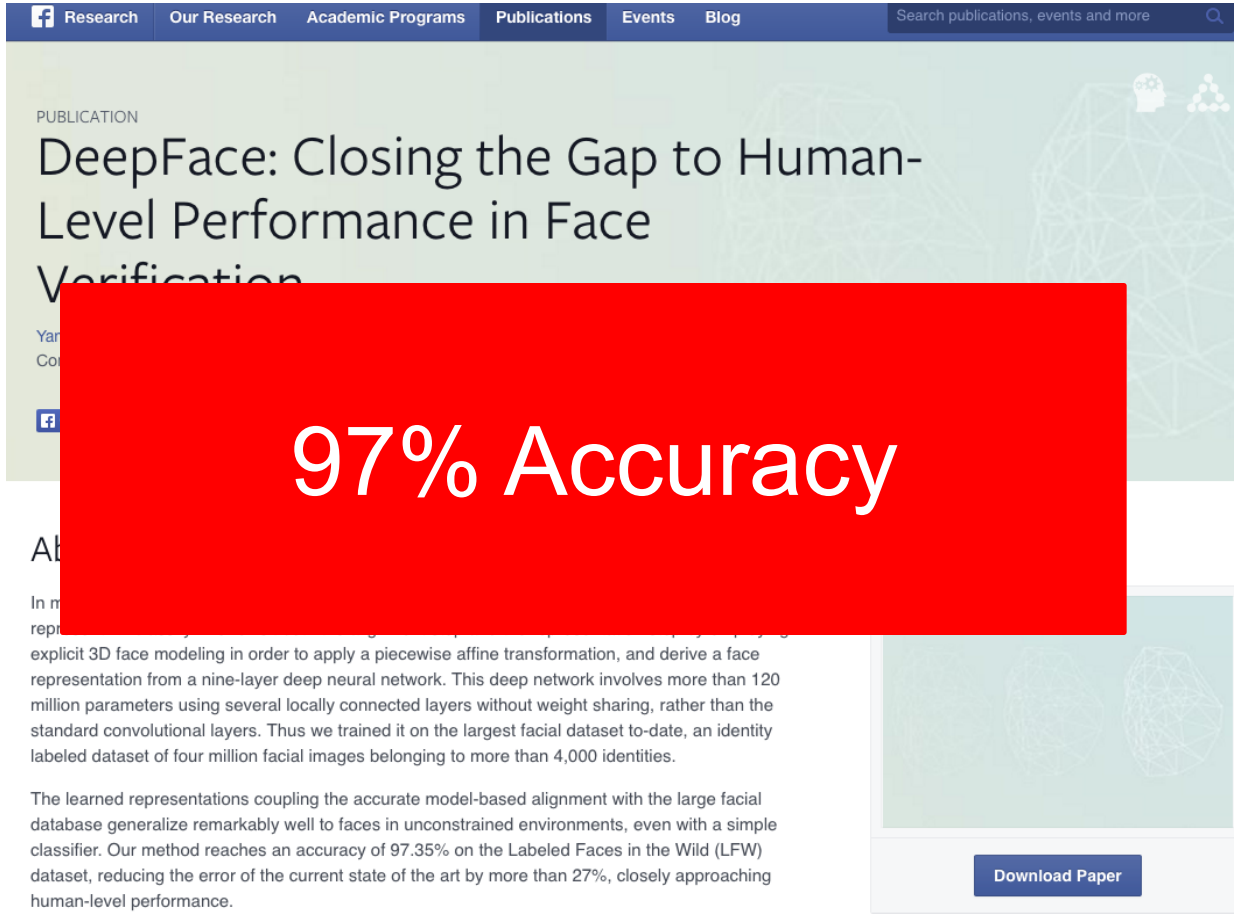
— Stephen Hawking —

Why?


IMAGENET Large Scale Visual Recognition Challenge
The Image Classification Challenge:
1,000 object classes
WON BY DEEP LEARNING
0.1
0
0.12
0.07
2010
2011
2012
2013
2014
ILSVRC year
Russakovsky et al. arXiv, 2014

Why?

MealoMi

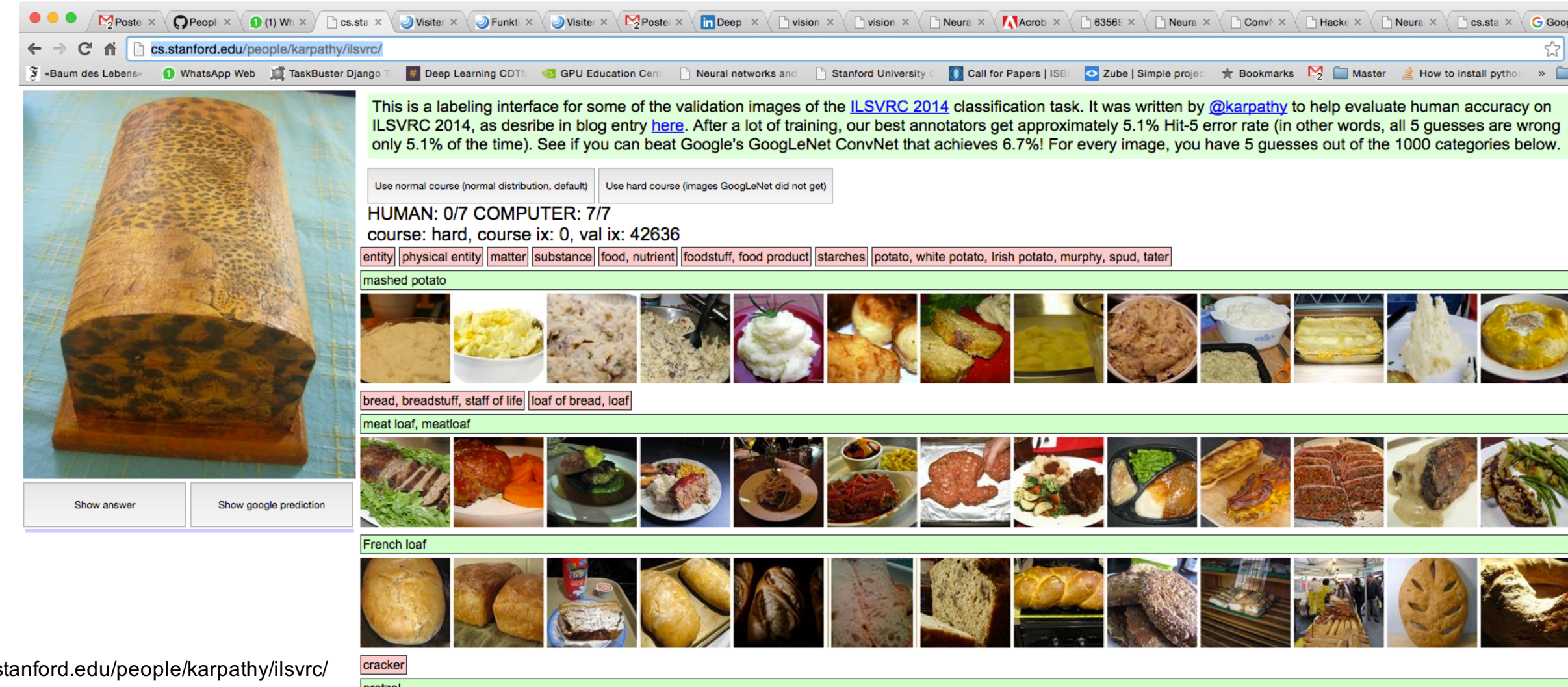

Research
Our Research
Academic Programs
Publications
Events
Blog
Search publications, events and more
PUBLICATION
DeepFace: Closing the Gap to Human-Level Performance in Face
97% Accuracy
explicit 3D face modeling in order to apply a piecewise affine transformation, and derive a face representation from a nine-layer deep neural network. This deep network involves more than 120 million parameters using several locally connected layers without weight sharing, rather than the standard convolutional layers. Thus we trained it on the largest facial dataset to-date, an identity labeled dataset of four million facial images belonging to more than 4,000 identities.
The learned representations coupling the accurate model-based alignment with the large facial database generalize remarkably well to faces in unconstrained environments, even with a simple classifier. Our method reaches an accuracy of 97.35% on the Labeled Faces in the Wild (LFW) dataset, reducing the error of the current state of the art by more than 27%, closely approaching human-level performance.
Download Paper

Who is who?

Mealoni

Test yourself

MealoMi



http://cs.stanford.edu/people/karpathy/ilsvrc/

Food recognition – Test Mealomi

MealoMi

AUTOMATIC FOOD DETECTION

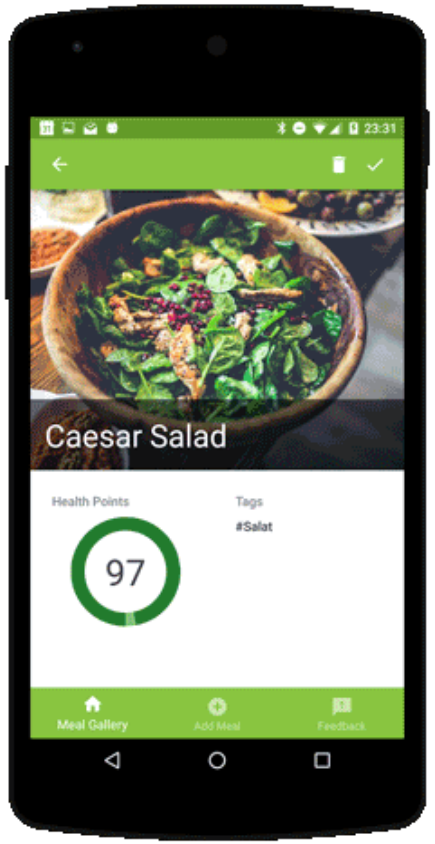
Detects your food automatically using the hottest algorithms based on deep learning.

NUTRITION INTAKE

Shows and stores your nutrition intake without typing it manually.

FOOD ANALYTICS

Shows stats and analytics about your personal food intake.



SIMPLE HEALTH SCORE

Calculates a simple health score to determine the healthiness of your food.

RECOMMENDATIONS

MealoMi learns your personal taste and recommends food, recipes and restaurants you like.

SHARING

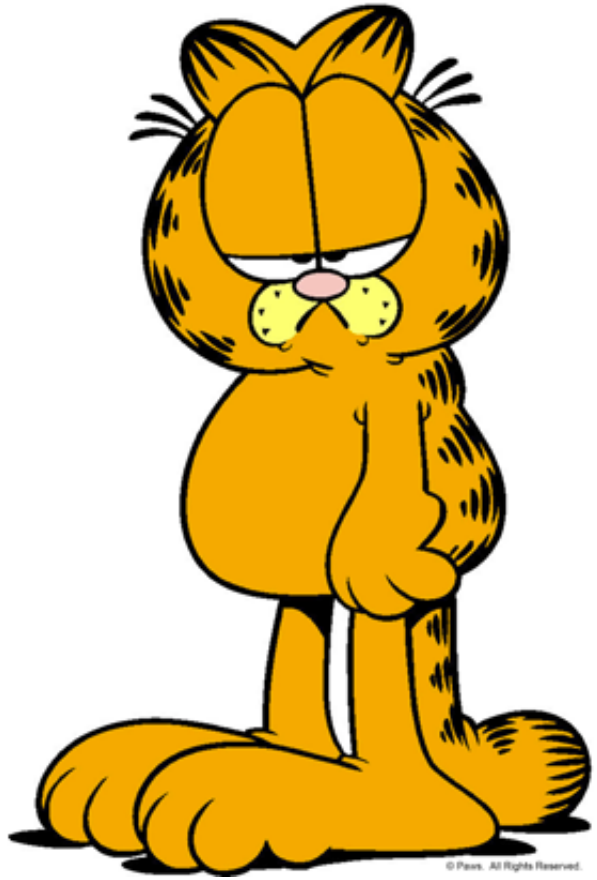
Share your food images and information with your friends.

Machine Learning – The Concept

MealoMi

VS

.

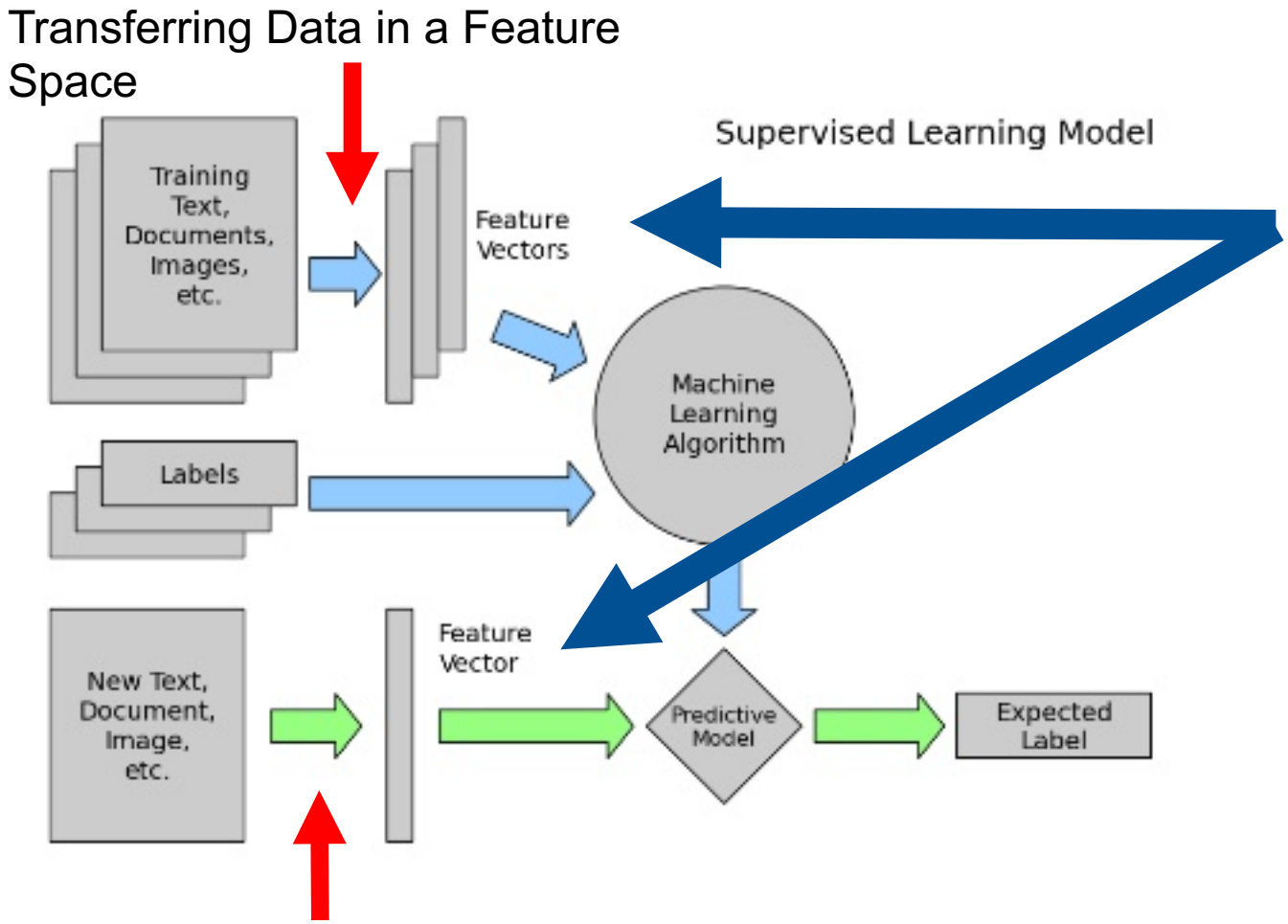


Classical Machine Learning Approach

MealoMi

Training Phase

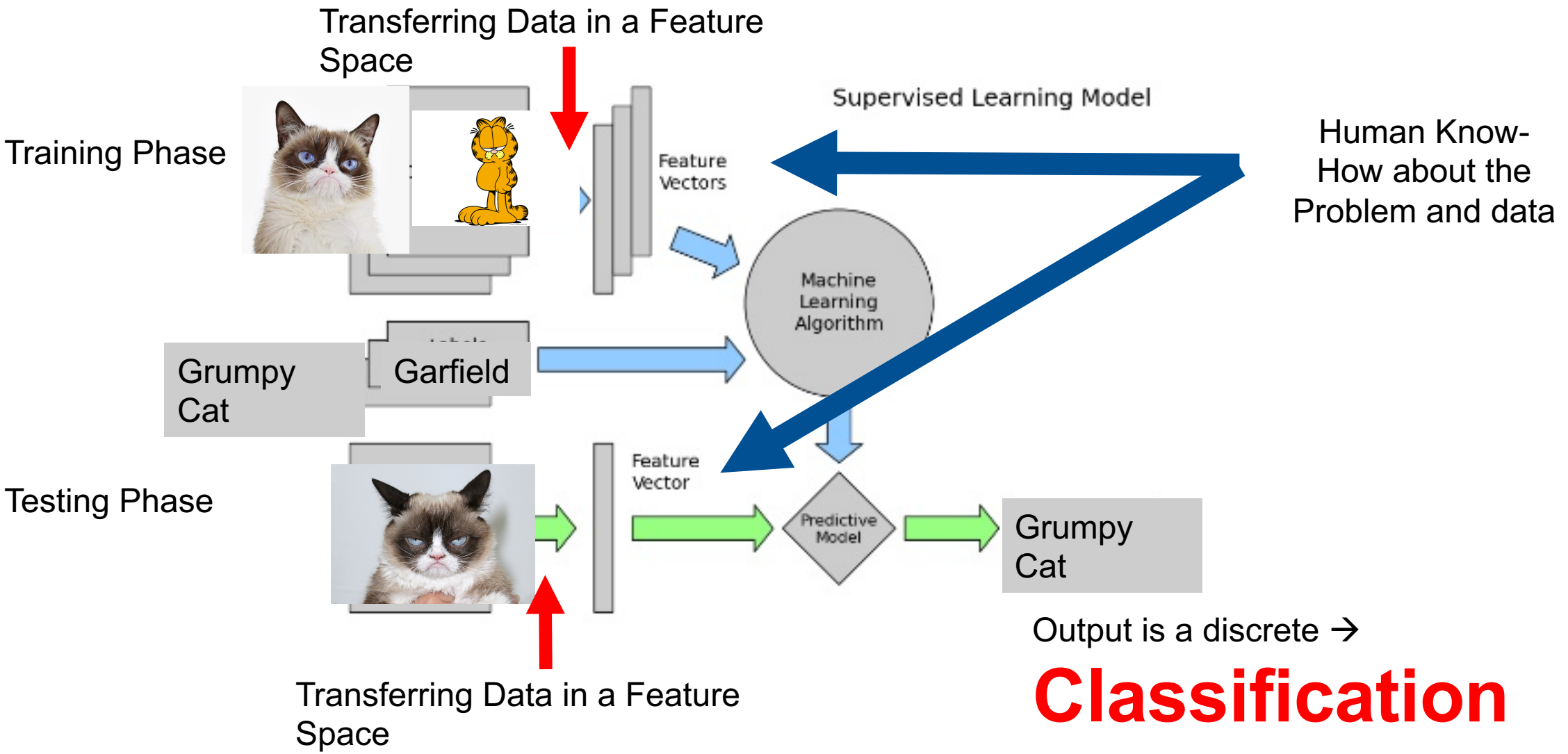
Testing Phase



Human Know-How about the Problem and data

Supervised Classification

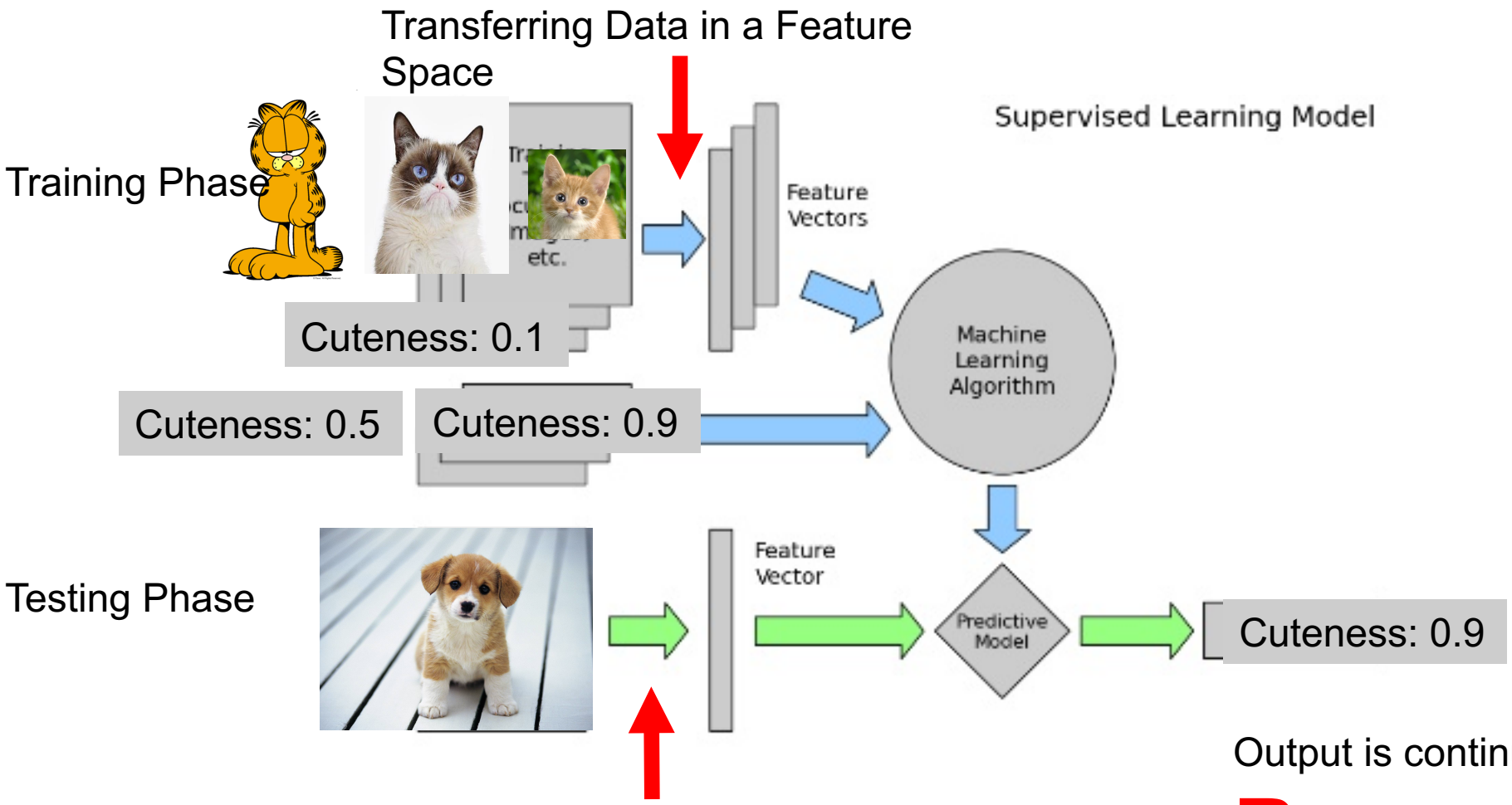
MealoMi



Output is a discrete →

Classification

Supervised Regression

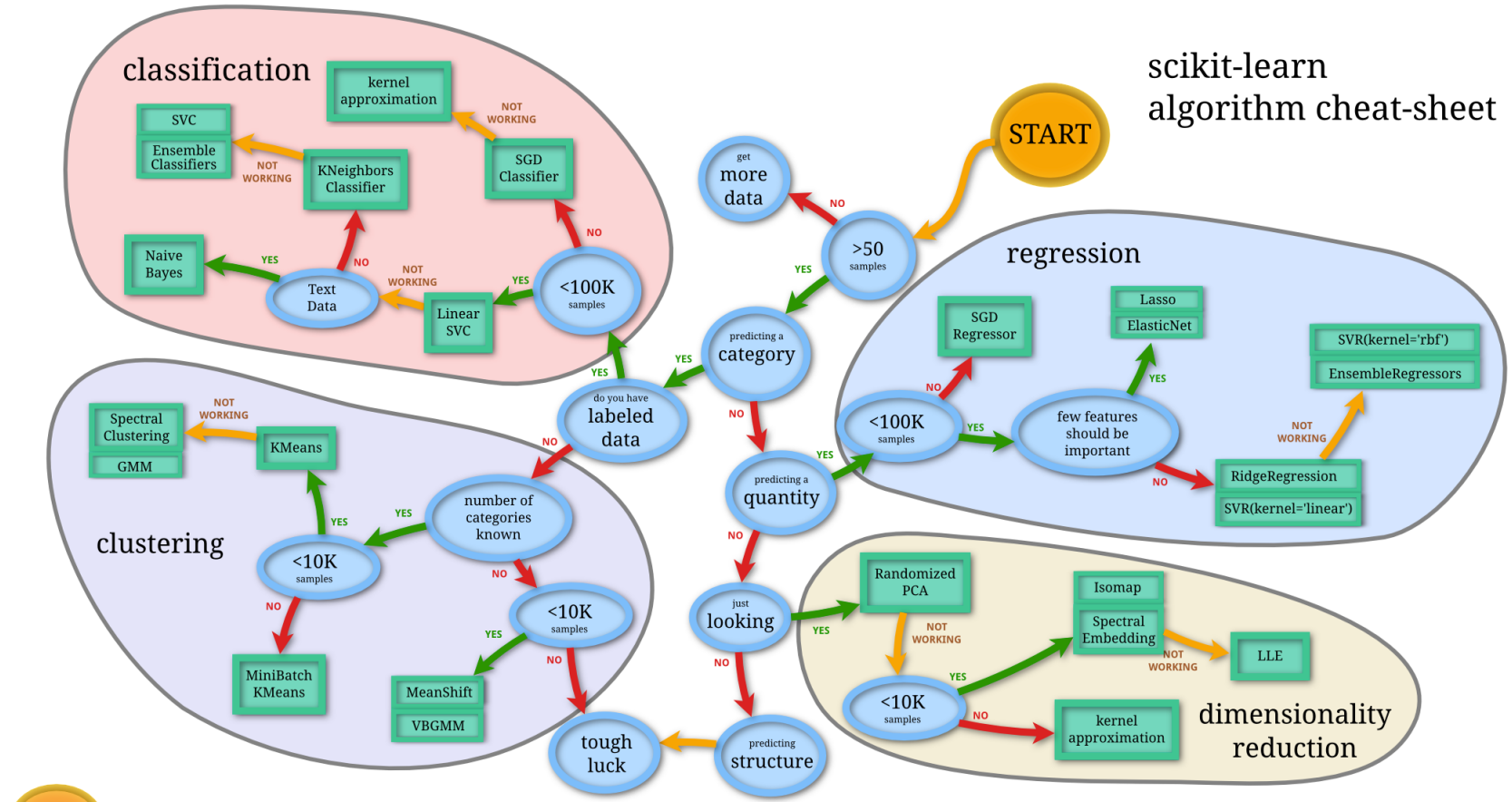


Output is continuous→

Regression

Machine Learning Cheat Sheet

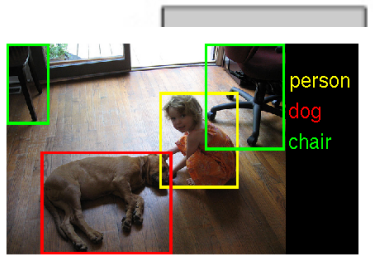
MealoMi



Supervised Regression for Object Detection

Transferring Data in a Feature Space

Training Phase



Supervised Learning Model

Feature Vectors

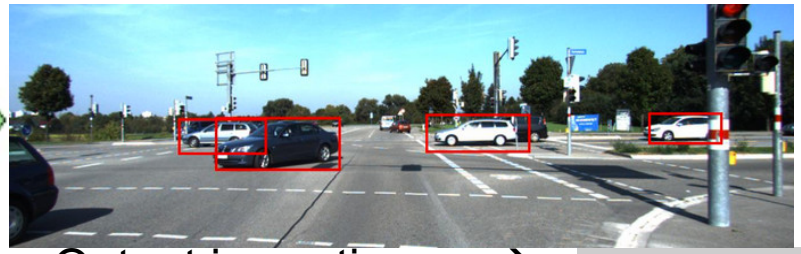
Machine Learning Algorithm

Person
Position x: 0.5
Position y: 0.3

Dog
Position x: 0.3
Position y: 0.5

Feature Vector

Predictive Model



Output is continuous→

Regressio

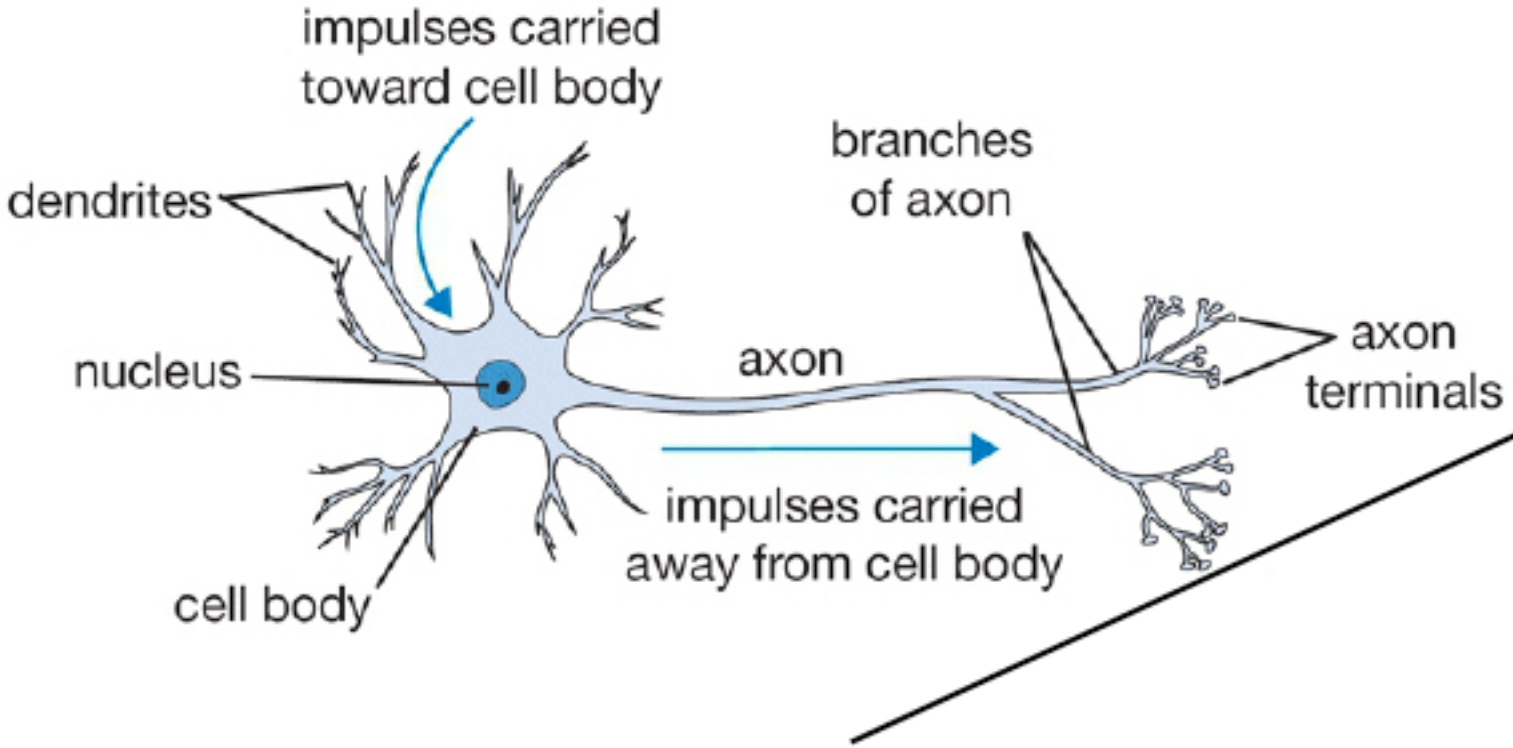
Car 1
Position x: 0.5
Position y: 0.3

Testing Phase

Transferring Data in a Feature Space

How does our brain work?

MealoMi



Input

Computation and Activation

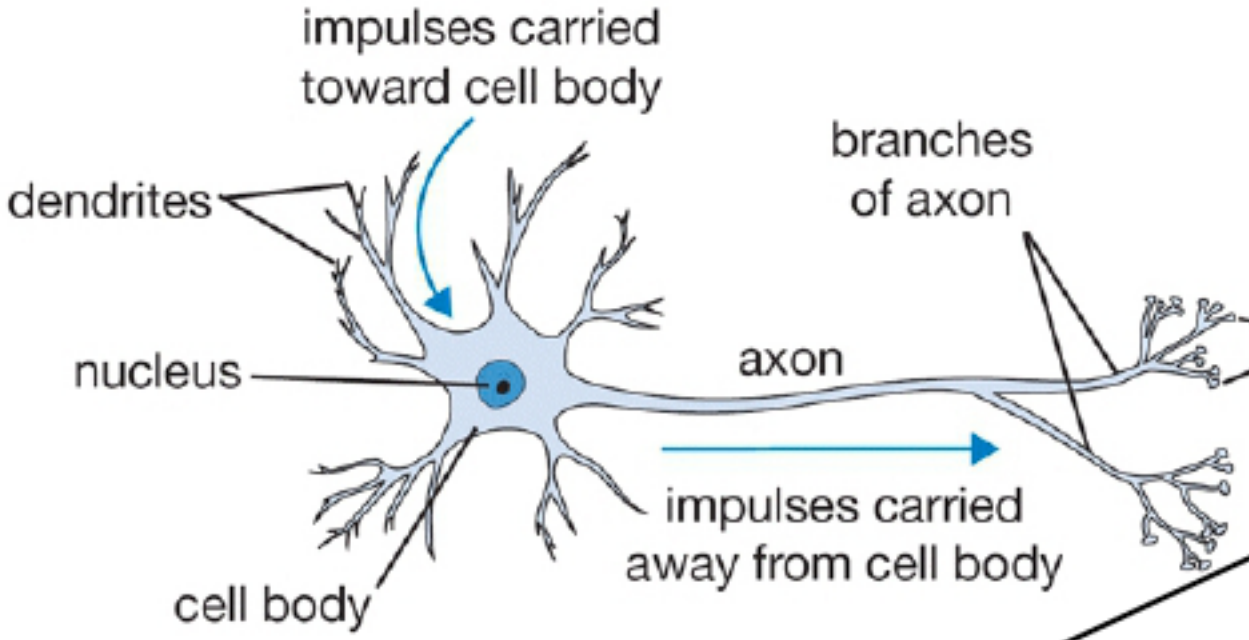
Output

The human brain has a huge number of synapses. Each of the 10^{11} (one hundred billion) neurons has on average 7,000 synaptic connections to other neurons. It has been estimated that the brain of a three-year-old child has about 10^{15} synapses (1 quadrillion). This number declines with age, stabilizing by adulthood. Estimates vary for an adult, ranging from 10^{14} to 5 x 10^{14} synapses (100 to 500 trillion).

Source: https://en.wikipedia.org/wiki/Neuron

How does our brain work?

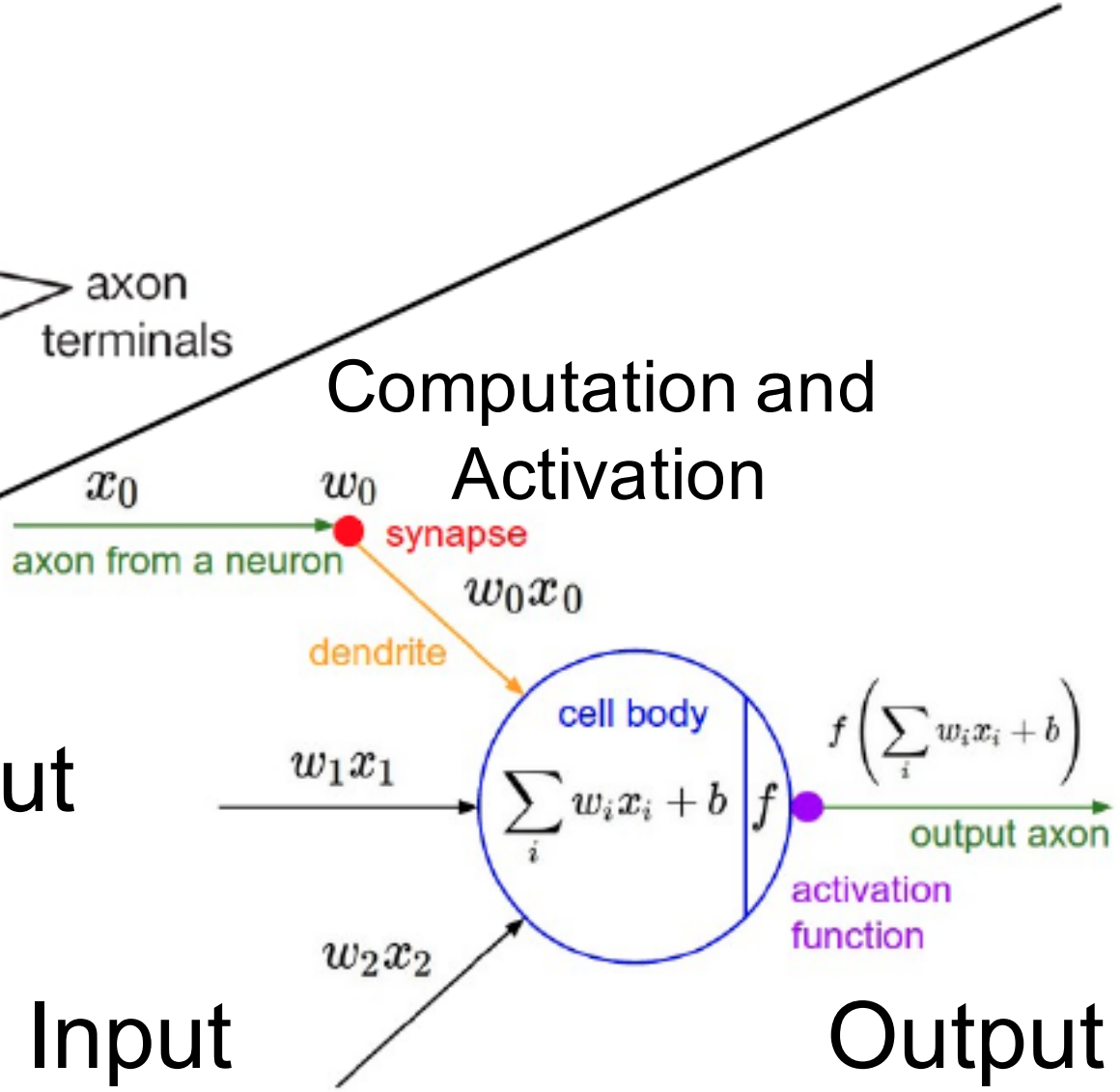
MealoMi



Input

Computation and
Activation

Output

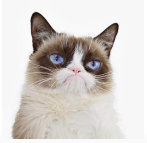


Input

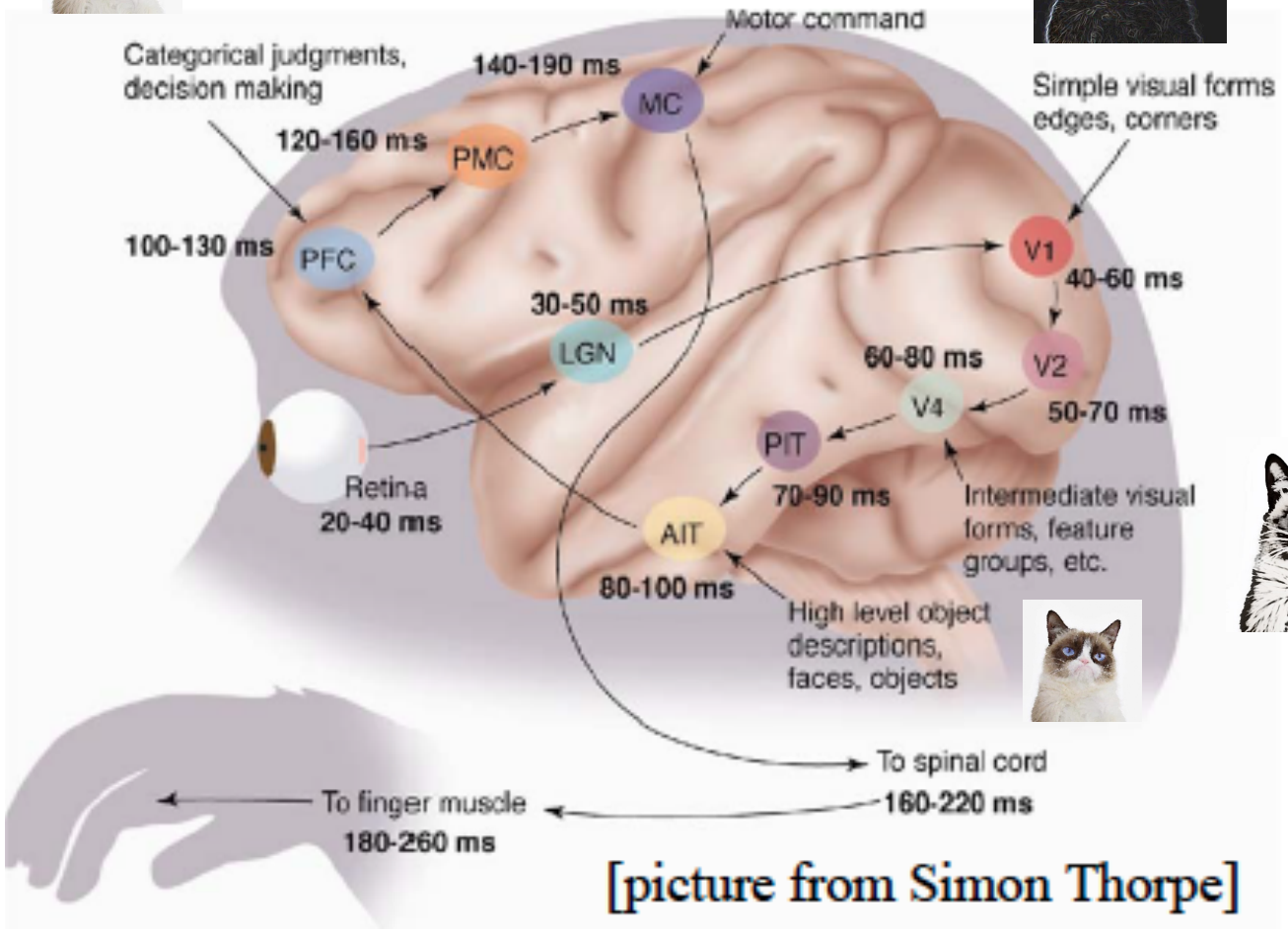
Output

How does our brain work?

MealoMi

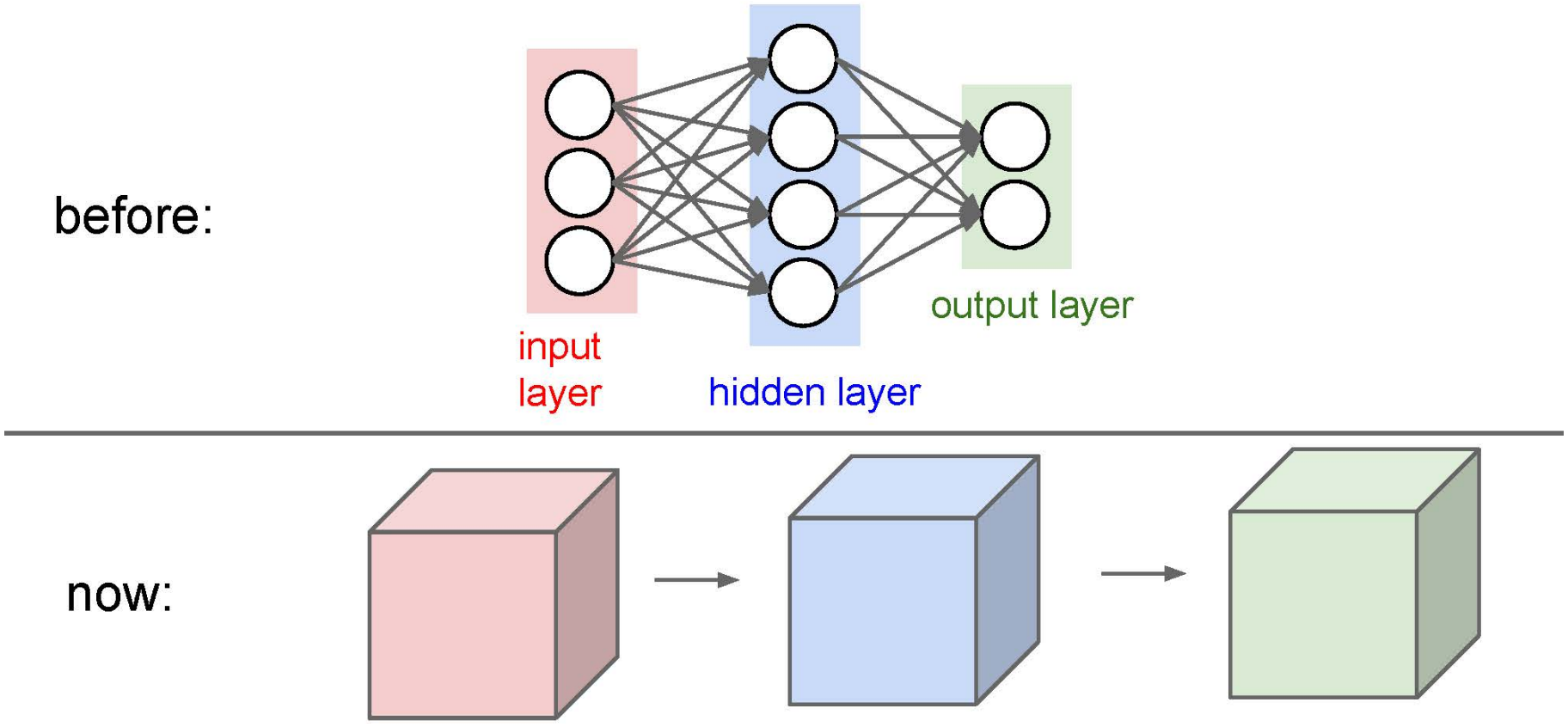


Grumpy cat → Touch



Neural Networks

MealoMi

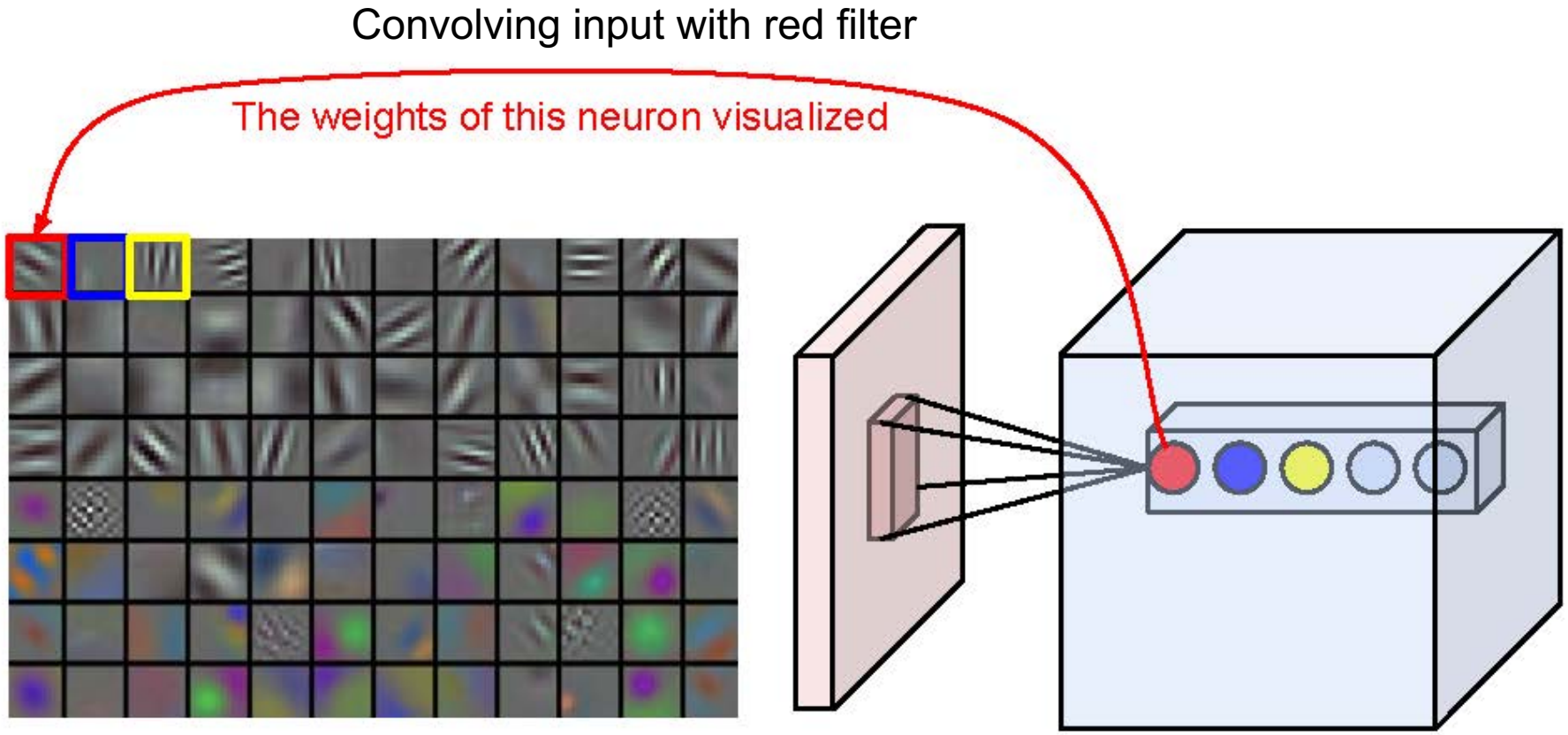


Credits: Fei-Fei and Karpathy http://vision.stanford.edu/teaching/cs23

Naming convention:

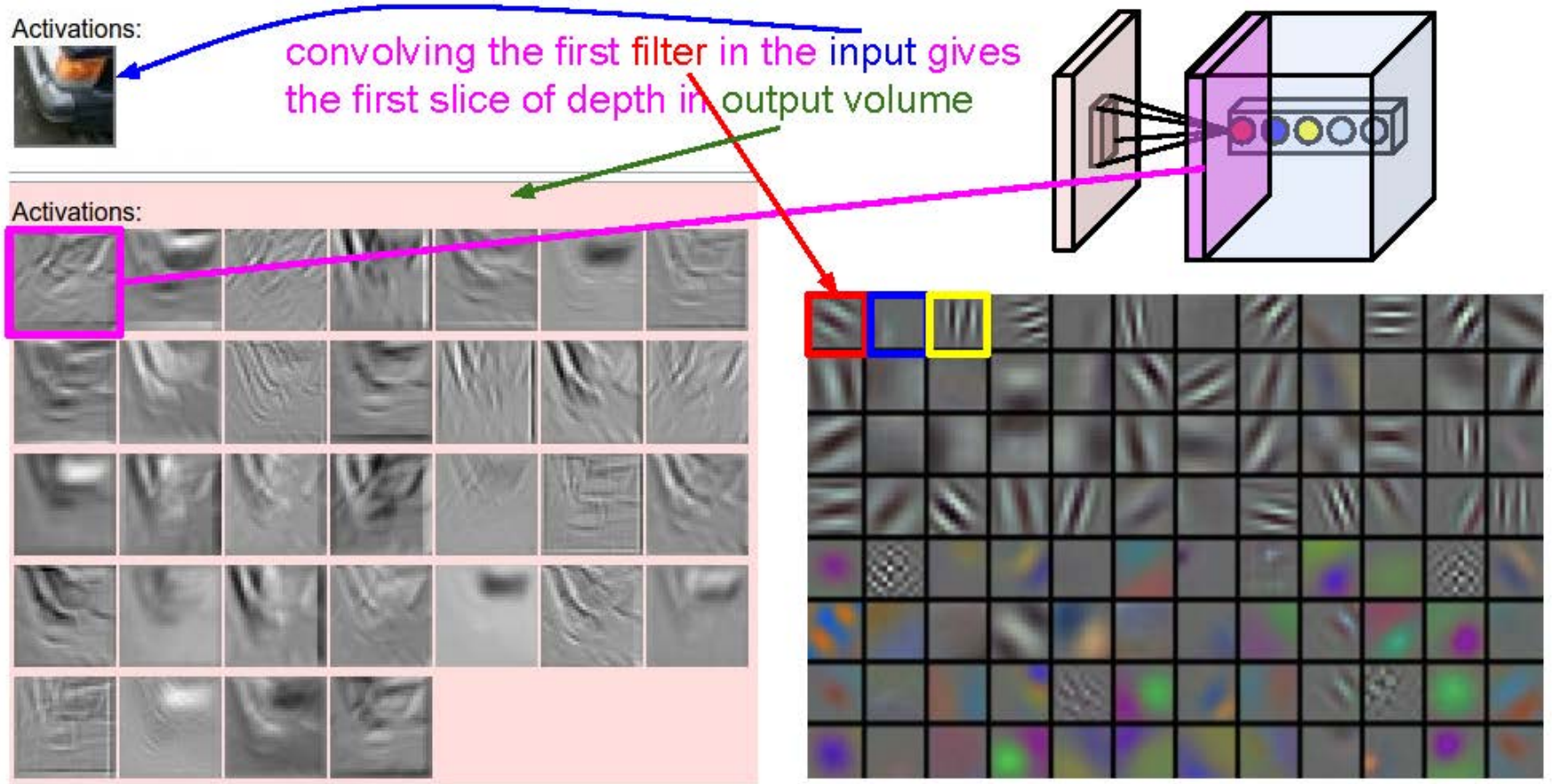
weights, filter, kernels are the same

Convolutional Neural Networks



Credits: Fei-Fei and Karpathy http://vision.stanford.edu/teaching/cs23

How the Convolution Works

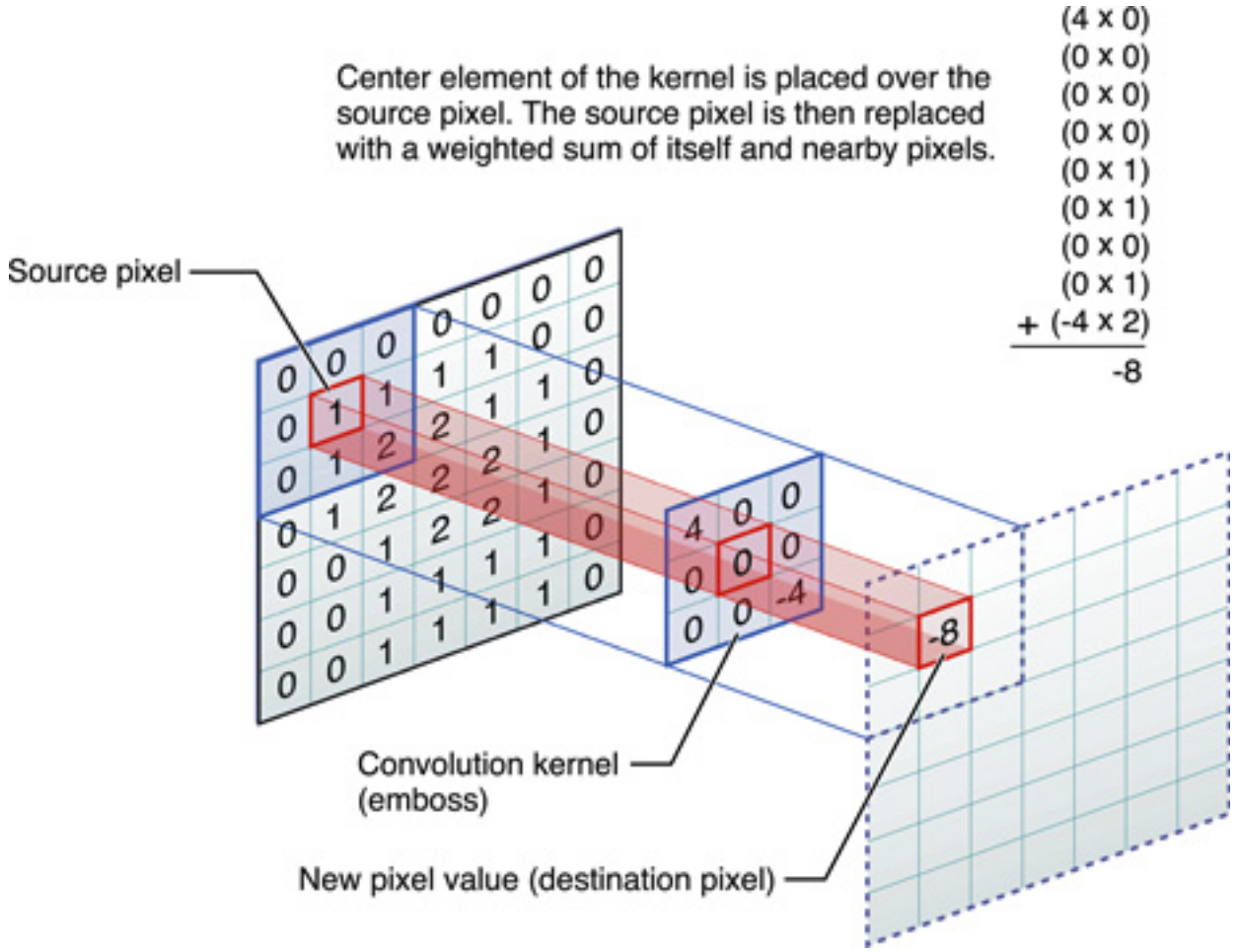


Credits: Fei-Fei and Karpathy http://vision.stanford.edu/teaching/cs23

Convolution

MealoMi

$$(f * g)(t) = \int f(\tau)\, g(t-\tau)\, d\tau$$

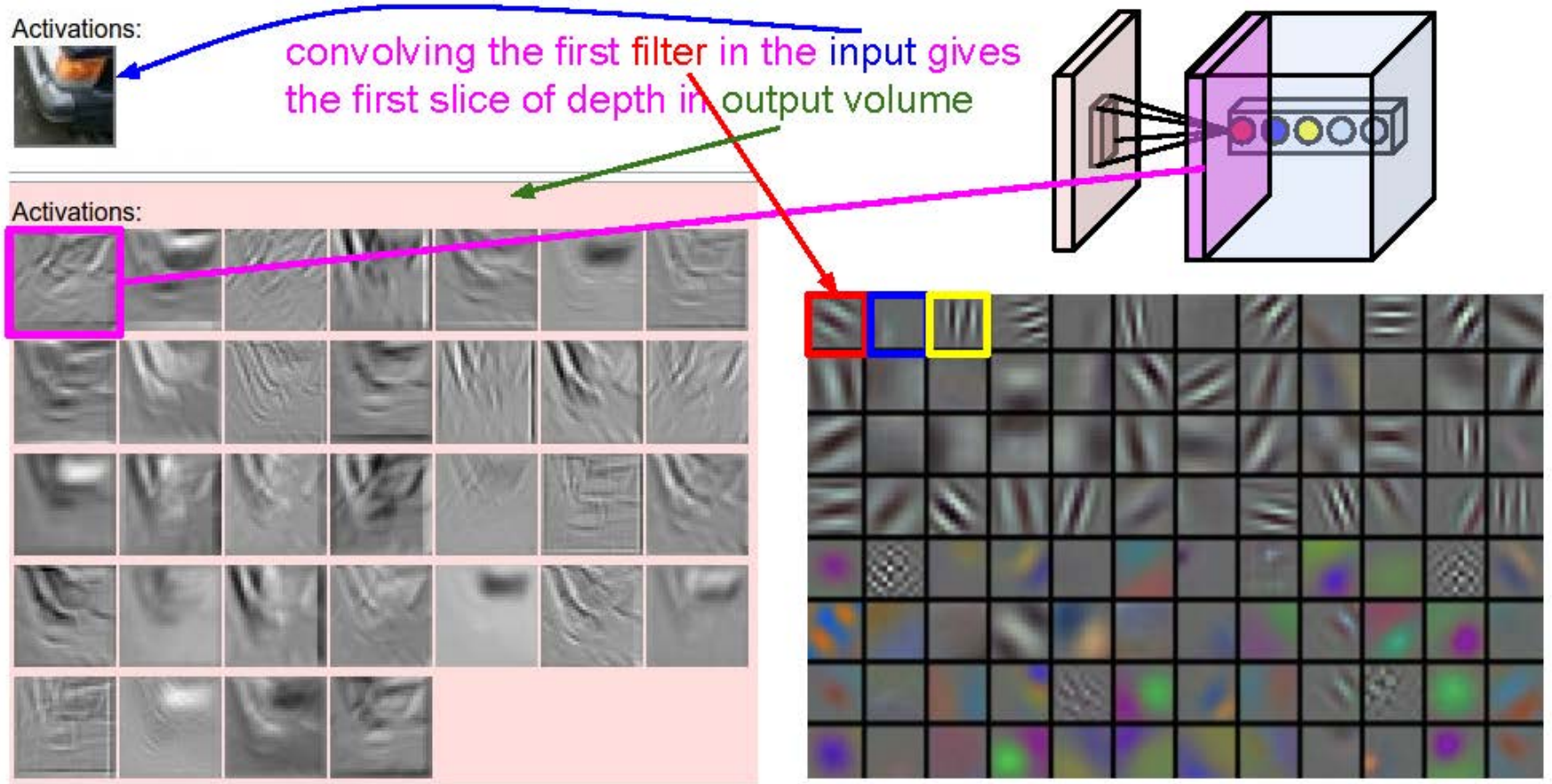


Convolution in Detail

MealoMi

http://cs231n.github.io/convolutional-networks/

How the Convolution Works

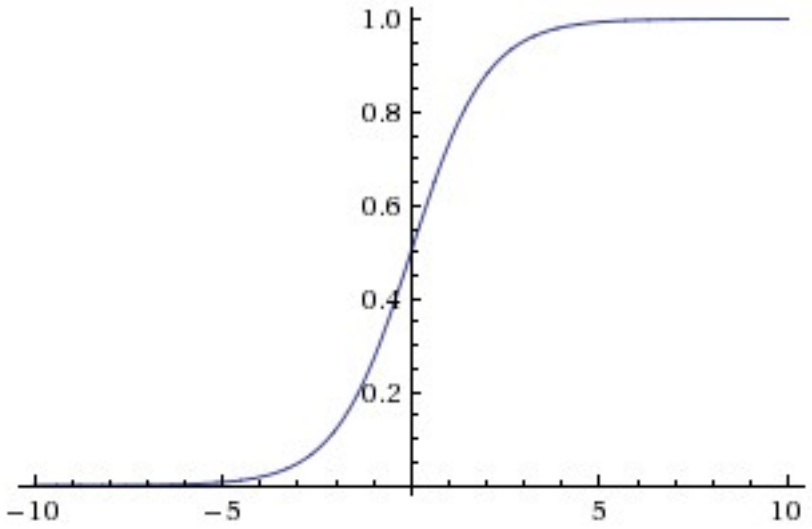


Credits: Fei-Fei and Karpathy http://vision.stanford.edu/teaching/cs23

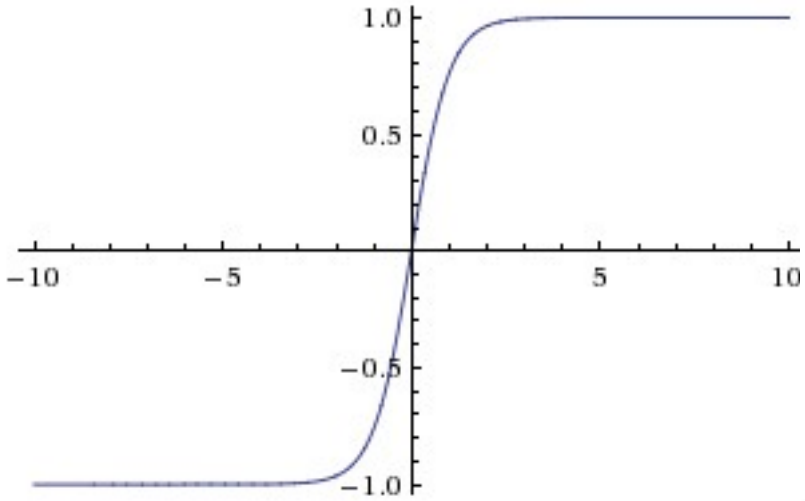
Activation

MealoMi

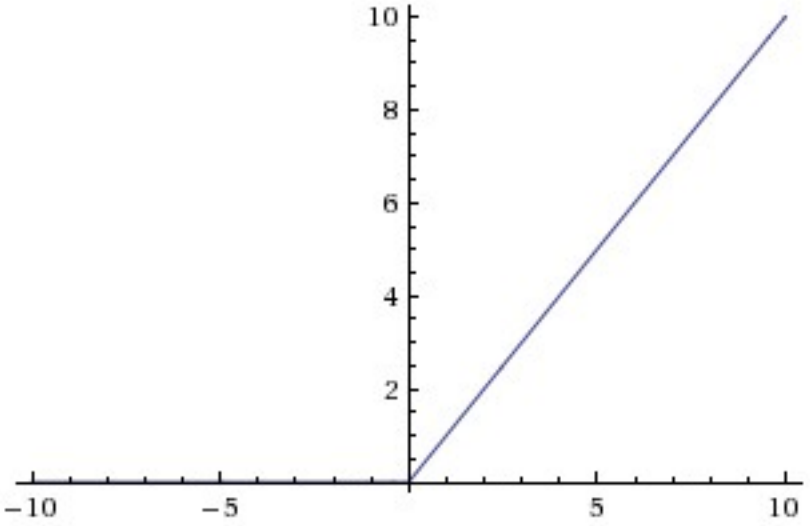
Idea: Decide whether the computation will be forwarded to the next layer or not



SIGMOID



TANH



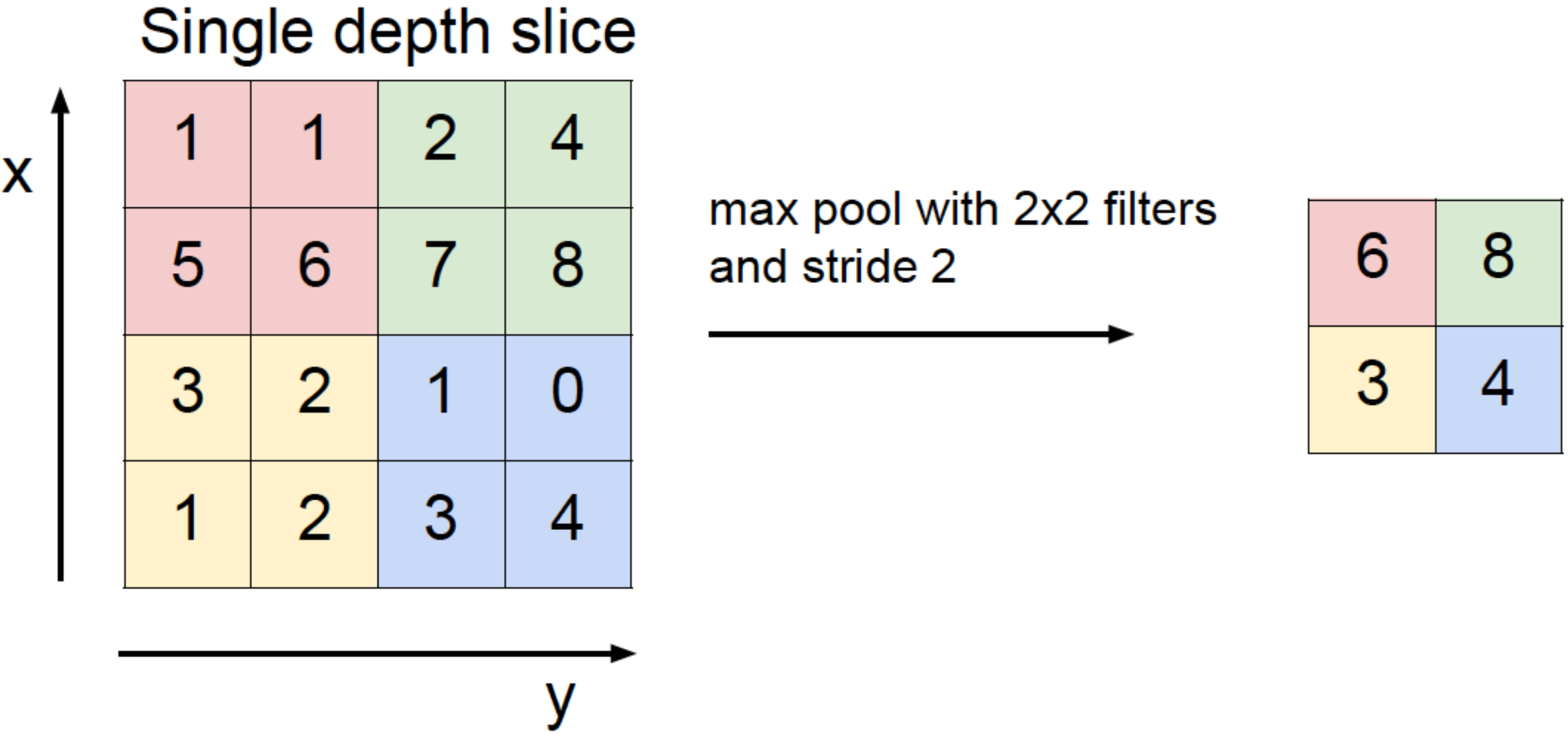
RELU

What is it good for?

Activations are activated if there is something useful in the image to solve the task.

http://cs.stanford.edu/people/karpathy/convnetjs/demo/cifar10.html

Max Pooling

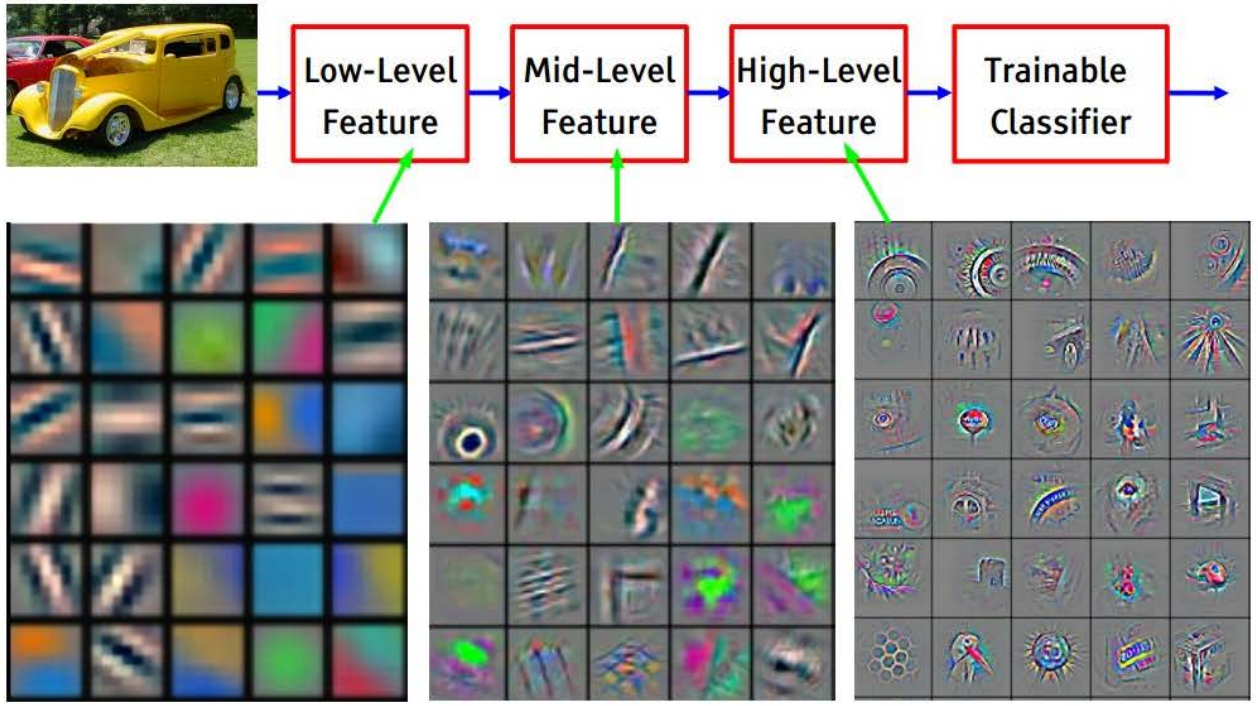

Single depth slice
x
y
1 1 2 4
5 6 7 8
3 2 1 0
1 2 3 4
max pool with 2x2 filters
and stride 2
6 8
3 4

Hierarchical Learning

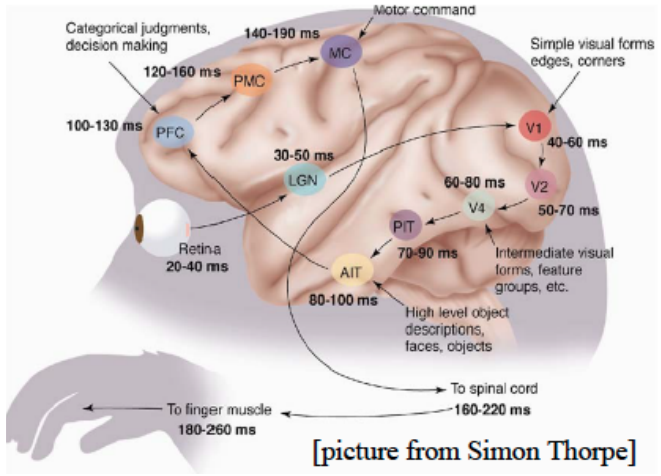
MealoMi

Fast-forward to today

[From recent Yann LeCun slides]



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



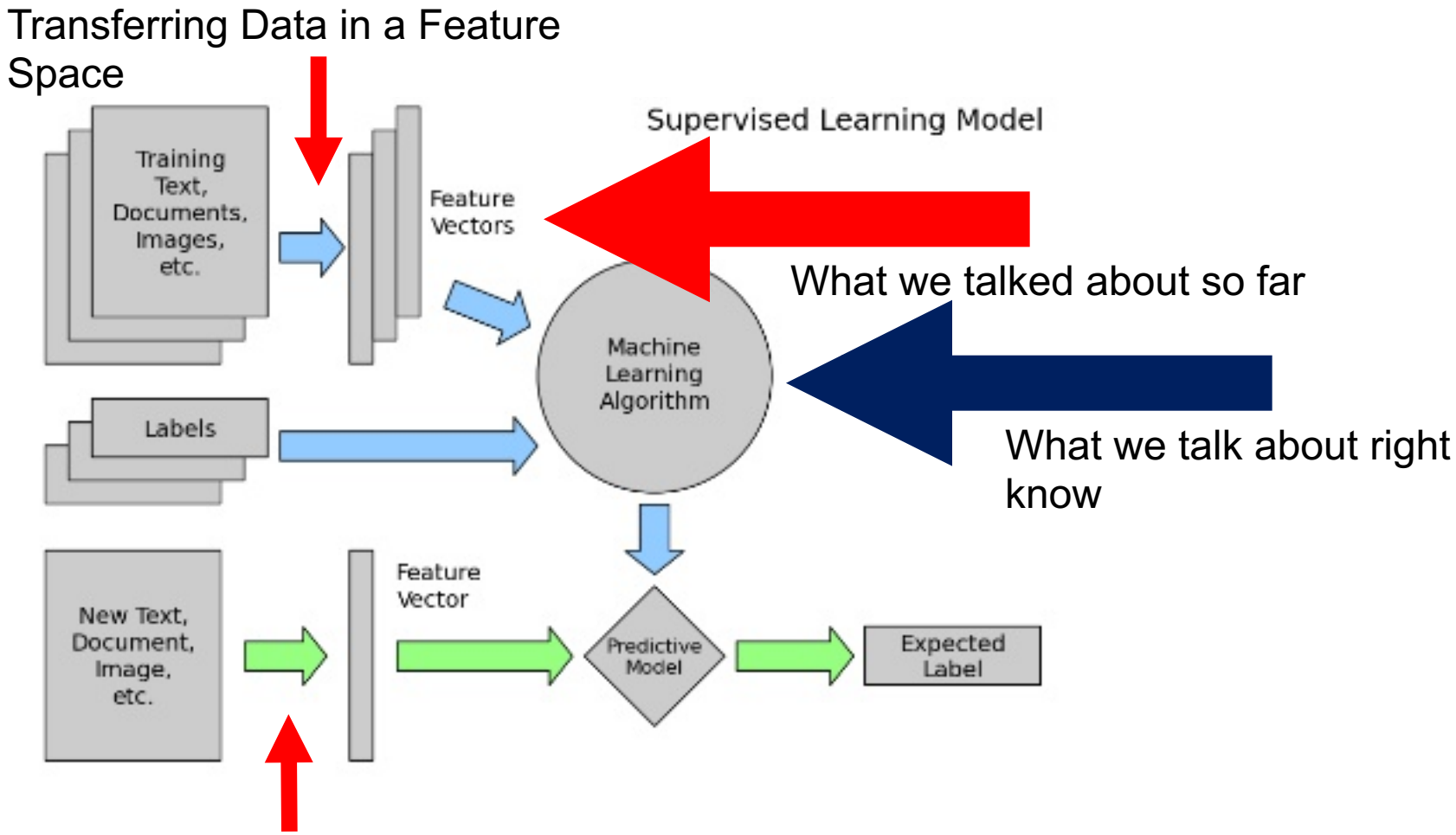
[picture from Simon Thorpe]

Classical Machine Learning Approach

Mealoni

Training Phase

Testing Phase



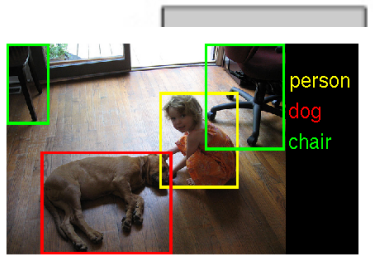
MealoMi

How to learn what is right or wrong?

Giving Feedback to our Algorithm

Transferring Data in a Feature Space

Training Phase



Supervised Learning Model

Feature Vectors

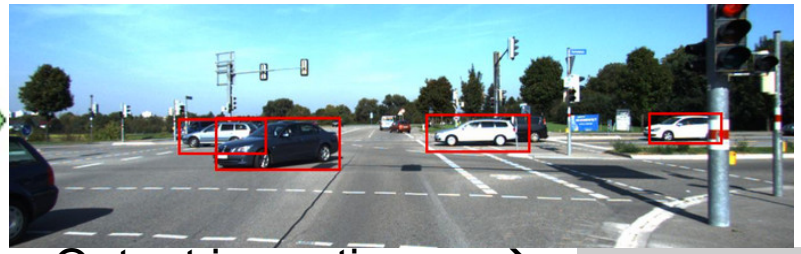
Machine Learning Algorithm

Person
Position x: 0.5
Position y: 0.3

Dog
Position x: 0.3
Position y: 0.5

Feature Vector

Predictive Model



Output is continuous→

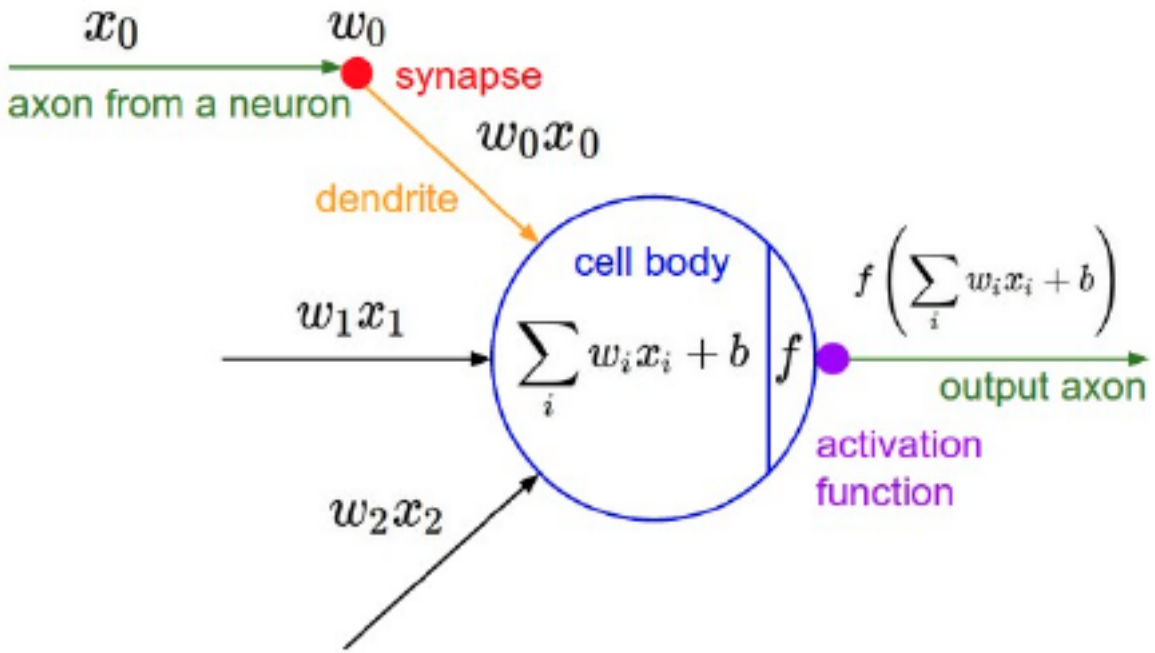
Regressio

Car 1
Position x: 0.5
Position y: 0.3

Testing Phase

Transferring Data in a Feature Space

Loss or Costfunction



1. Score function

$$f(x_i, W, b) = Wx_i + b$$

2. Two loss functions

$$L_i = -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right)$$

$$L_i = \sum_{j \neq y_i} \max(0, f(x_i, W)_j - f(x_i, W)_{y_i} + \Delta)$$

Goal: Minimize the Loss given x_i and y_i → Find weights/filters that minimize the Loss

Gradient Descent and Backpropagation

MealoMi

Idea: Minimize Loss/Costfunction given the prediction of my current network $f(x_i w_i+b)$ and my true label y_i.

→This Problem is not analytical solvable ☹
→Numerical method to do this is called Stochastic Gradient Decent
http://neuralnetworksanddeeplearning.com/chap3.html
The algorithm to calculate the loss of a certain layer is called Backpropagation.
http://neuralnetworksanddeeplearning.com/chap2.html

What we not covered yet

MealoMi

Padding:

Slide 45: http://vision.stanford.edu/teaching/cs231n/slides/lecture7.pdf

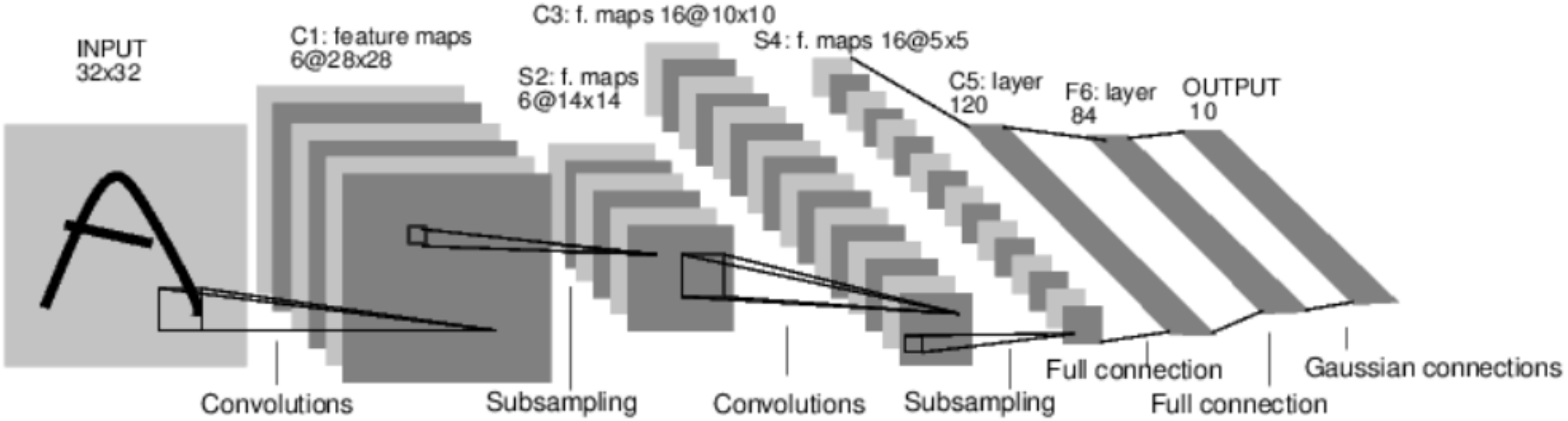
Training Hyperparameter

http://neuralnetworksanddeeplearning.com/chap3.html

Deep Network in action

MetaMind

http://cs.stanford.edu/people/karpathy/convnetjs/demo/mnist.html



More Examples at:
http://cs.stanford.edu/people/karpathy/convnetjs/

[LeNet-5, LeCun 1980]

Finetuning of a Deep Network

MealoMi

Training a deep network with a lot of layers takes a lot of and is sometimes not possible, because you have not enough data.

Idea: Do not start from scratch. Take the learned Results (models) from other people and learn further on your problem.

Advantage: + save time and train even with less data
Resources: http://caffe.berkeleyvision.org/model_zoo.html
https://github.com/BVLC/caffe/wiki/Model-Zoo

More Resources

MealoMi

http://neuralnetworksanddeeplearning.com/chap1.html
http://deeplearning.net/
http://caffe.berkeleyvision.org/
http://vision.stanford.edu/teaching/cs231n/syllabus.html
http://goodfeli.github.io/dlbook/

Super Resources:
https://nvidia.qwiklab.com/learning_paths/10/lab_catalogue

MealoMi

Start with Coding 2

https://www.youtube.com/watch?v=bHvf7Tagt18