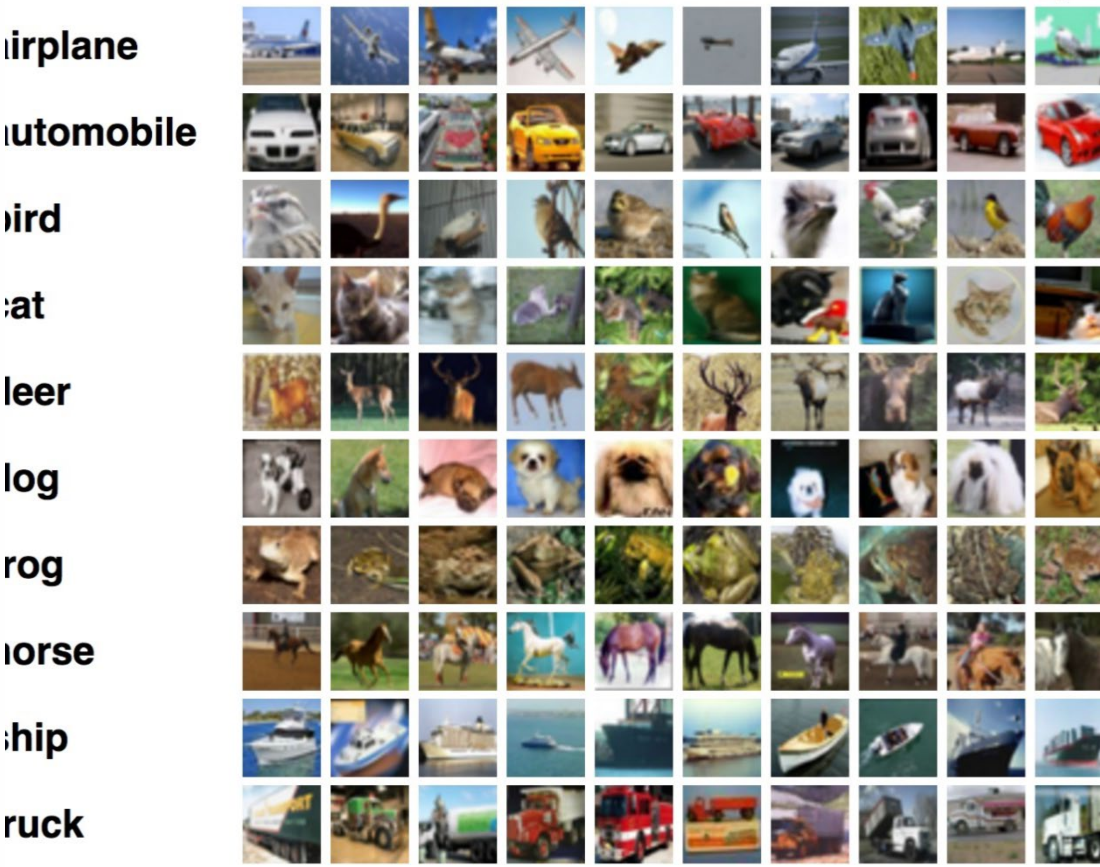
DEEP LEARNING:
IMAGE CLASSIFICATION



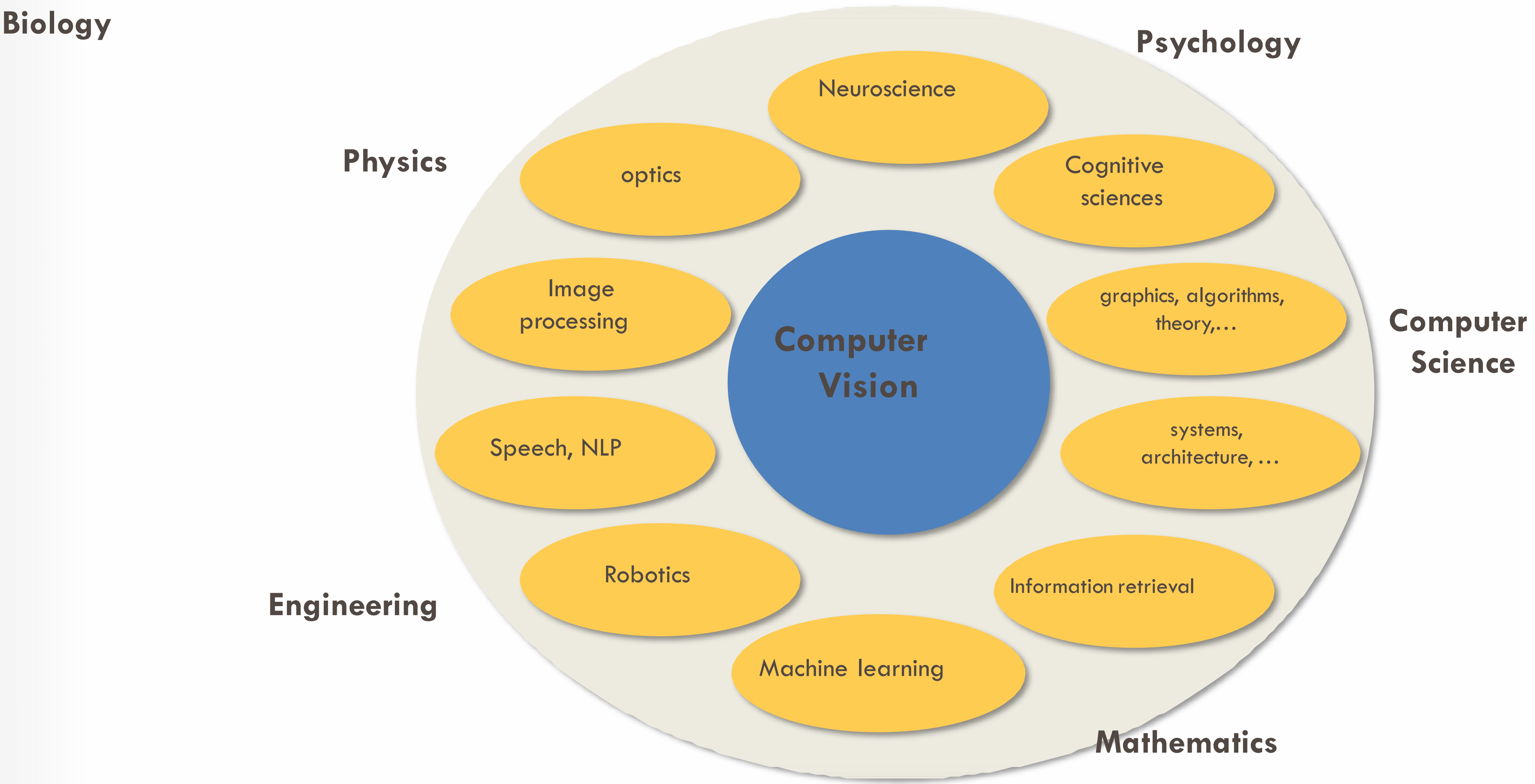




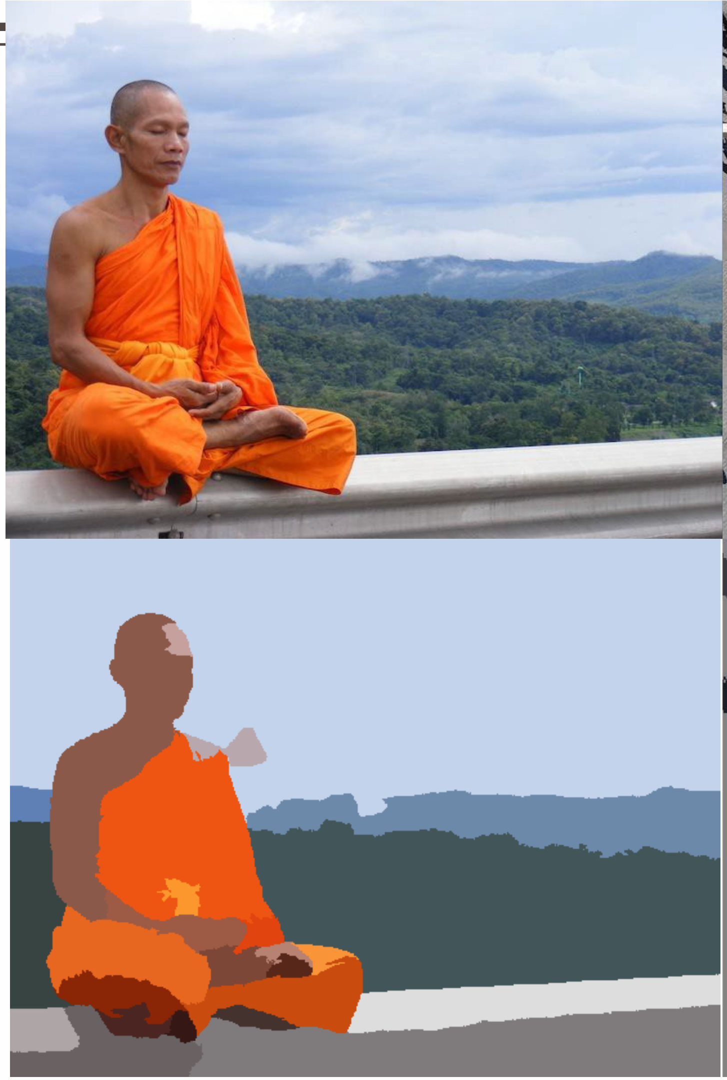


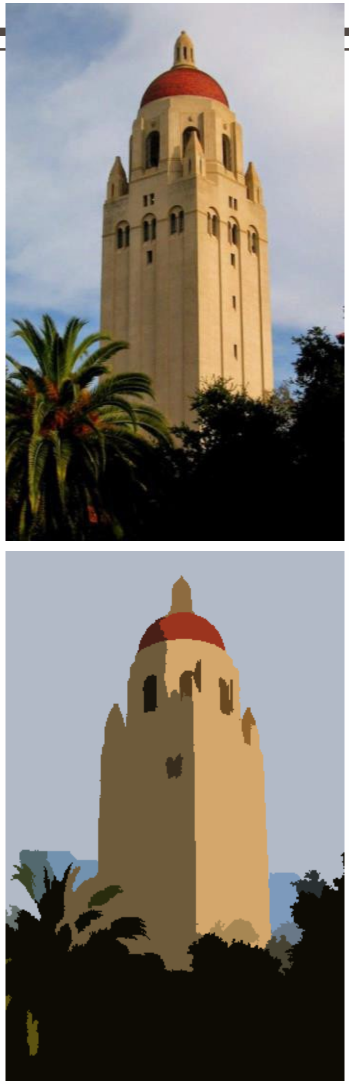






Biology
Psychology
Neuroscience
Physics
optics
Cognitive sciences
Image processing
Computer Vision
graphics, algorithms, theory,…
Computer Science
Speech, NLP
systems, architecture, …
Robotics
Information retrieval
Engineering
Machine learning
Mathematics

Normalized Cut (Shi & Malik, 1997)





Face Detection, Viola & Jones, 2001

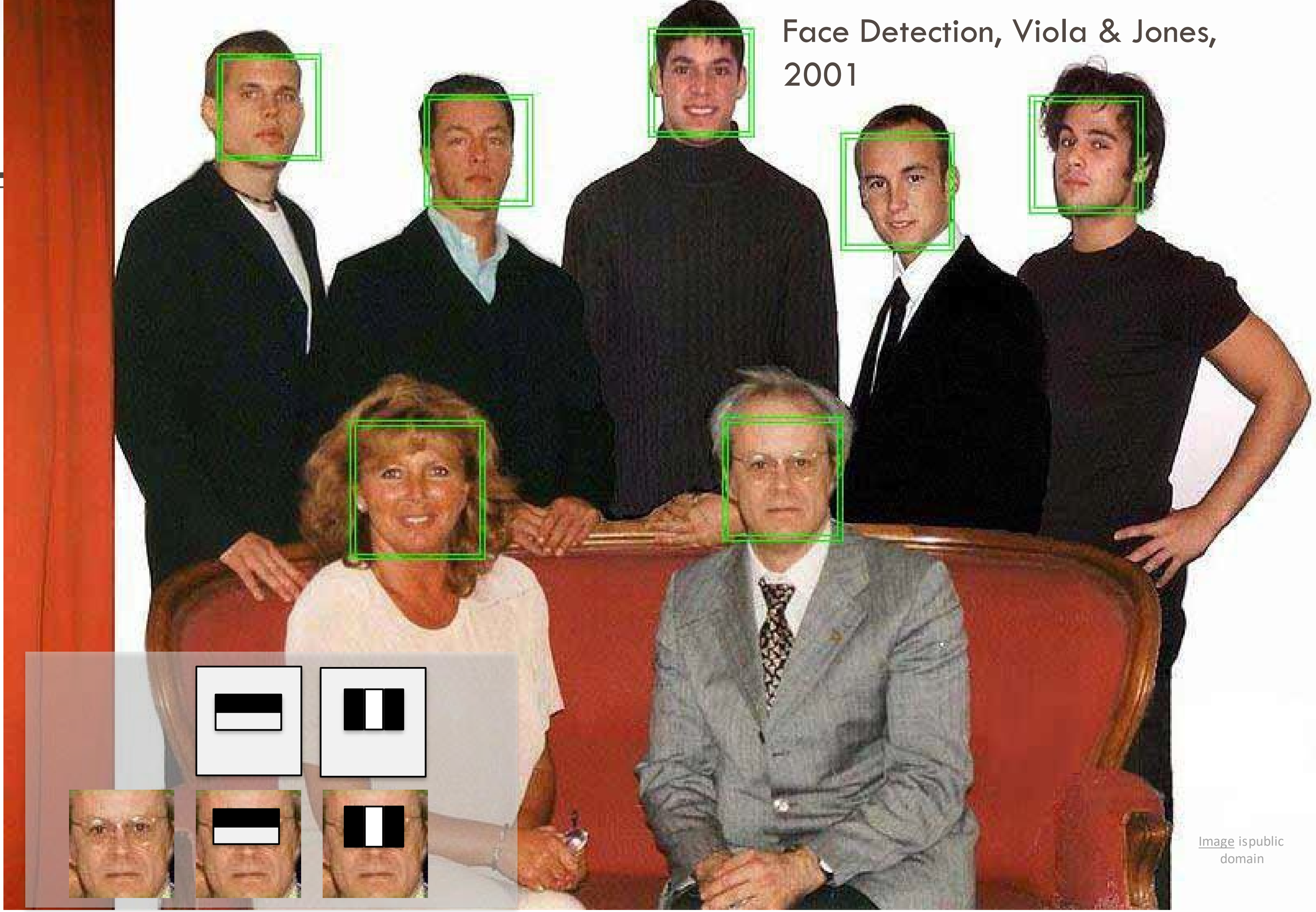


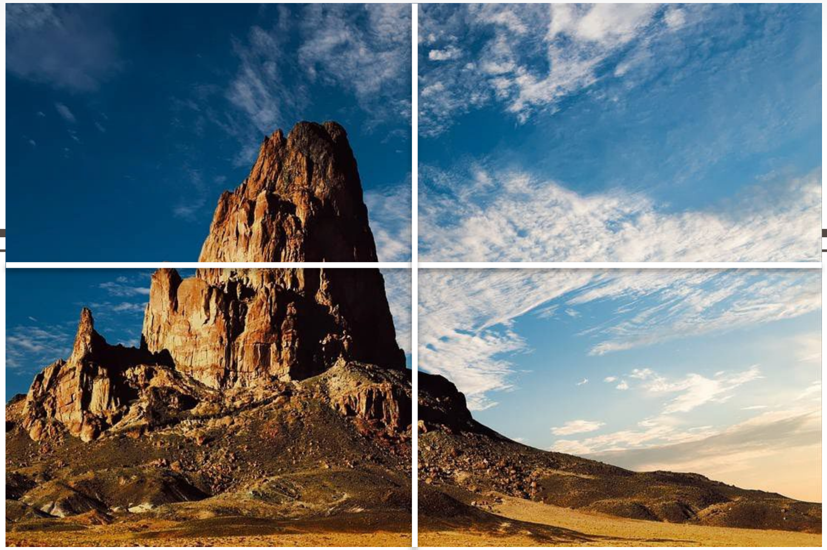
Image is public domain

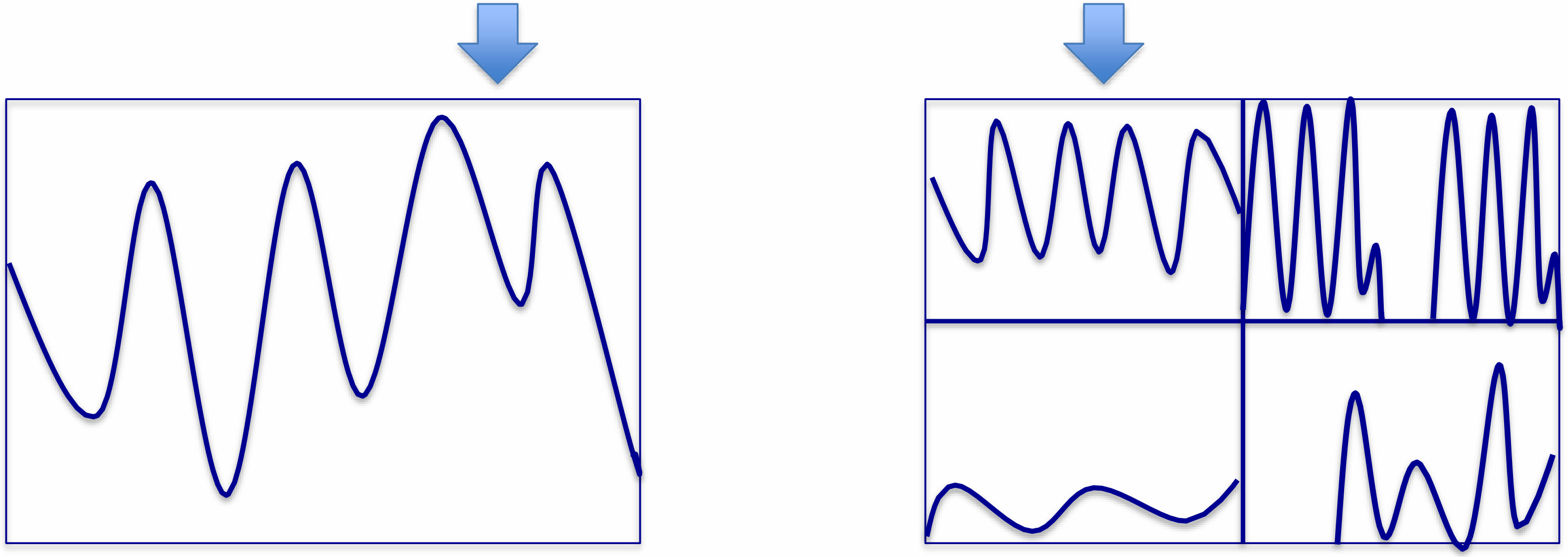
Local Feature and Matching

Image is public domain

Image is public domain

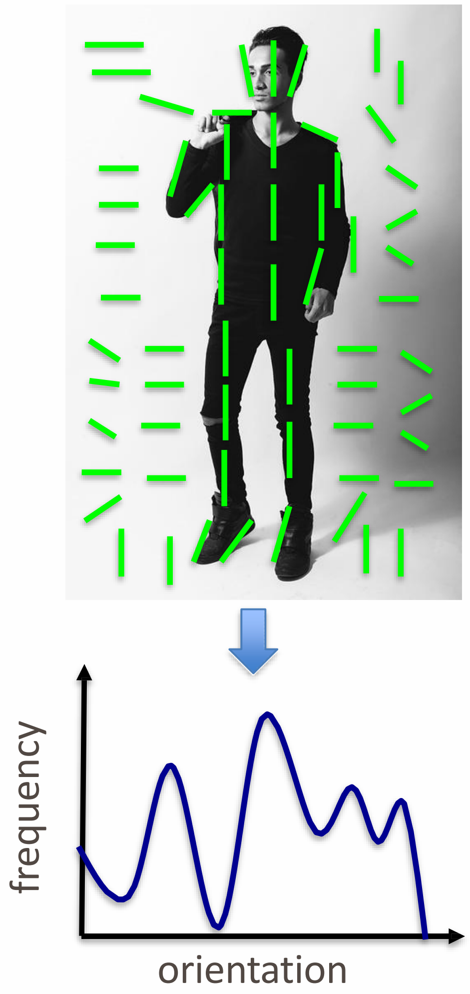
"SIFT" & Object Recognition, David Lowe, 1999



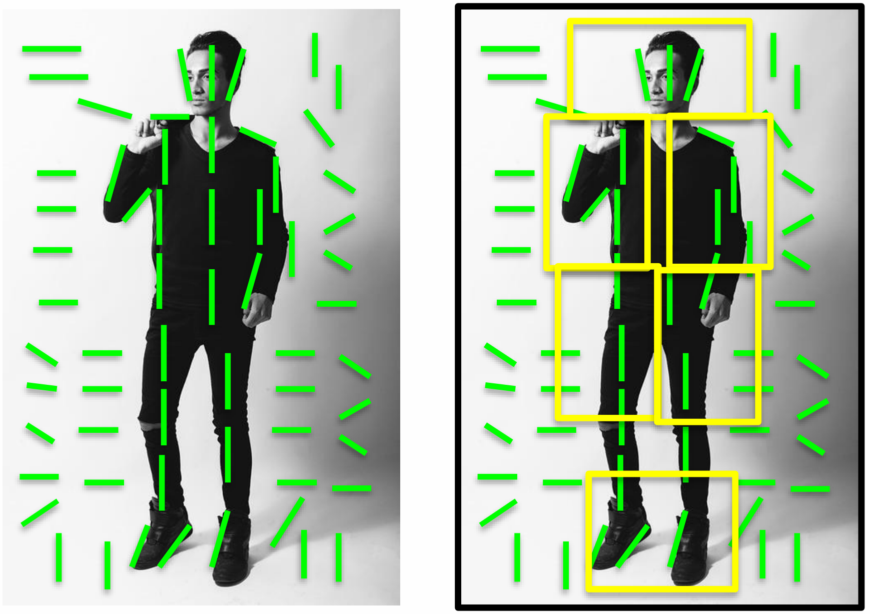


Level 0 Level 1

Spatial Pyramid Matching, Lazebnik, Schmid & Ponce, 2006



Histogram of Gradients (HoG)
Dalal & Triggs, 2005



Deformable Part Model
Felzenswalb, McAllester, Ramanan, 2009

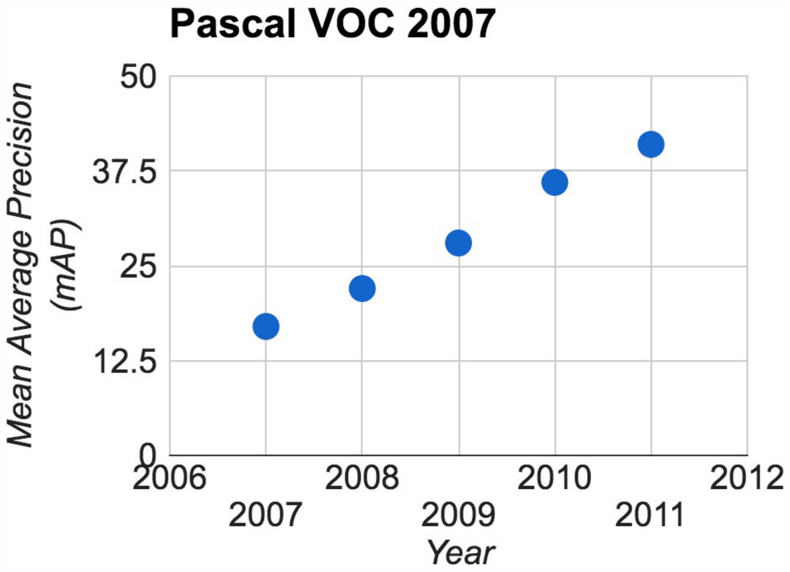
PASCAL Visual Object Challenge (20 object categories)

[Everingham et al. 2006-2012]

Image is CC0 1.0 public domain

Image is CC0 1.0 public domain

Image is CC0 1.0 public domain



IMAGENET Large Scale Visual Recognition Challenge

The Image Classification Challenge: 1,000 object classes

1,431,167 images

Output:
Scale

<u>Steel drum</u>
Drumstick
Mud turtle

✓

Output:
Scale

Giant panda
Drumstick
Mud turtle

✗

Russakovsky et al. IJCV 2015

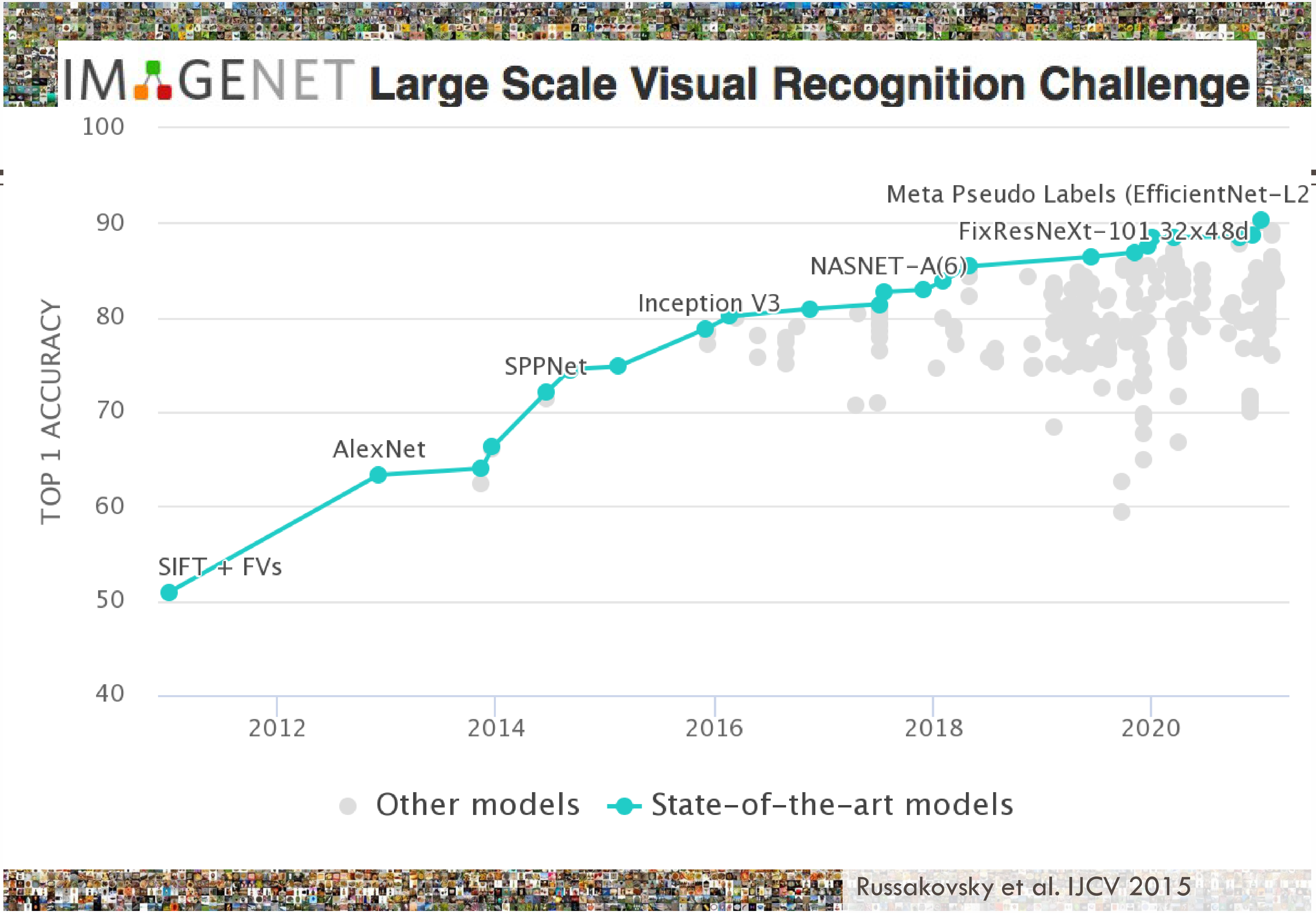

IMAGENET Large Scale Visual Recognition Challenge
100
90
80
70
60
50
40
TOP 1 ACCURACY
Meta Pseudo Labels (EfficientNet-L2
FixResNeXt-101 32x48d
NASNET-A(6)
Inception V3
SPPNet
AlexNet
SIFT + FVs
2012
2014
2016
2018
2020
Other models
State-of-the-art models
Russakovsky et al. IJCV 2015

Image by US Army is licensed under CC BY 2.0

Image is CC0 1.0 public domain

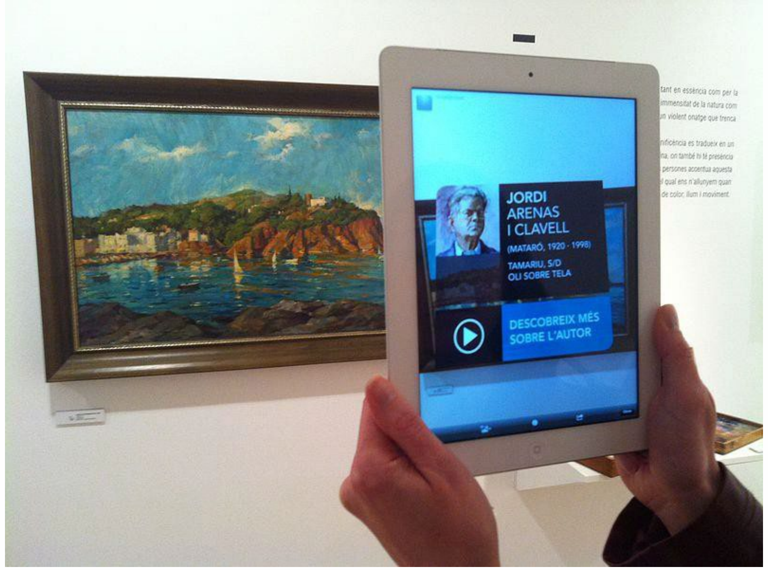
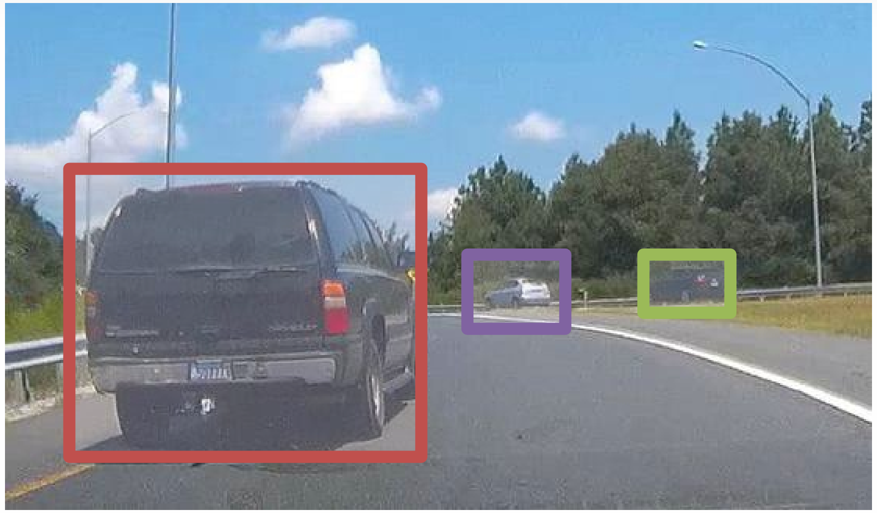


Image by Kippelboy is licensed under CC BY-SA 3.0

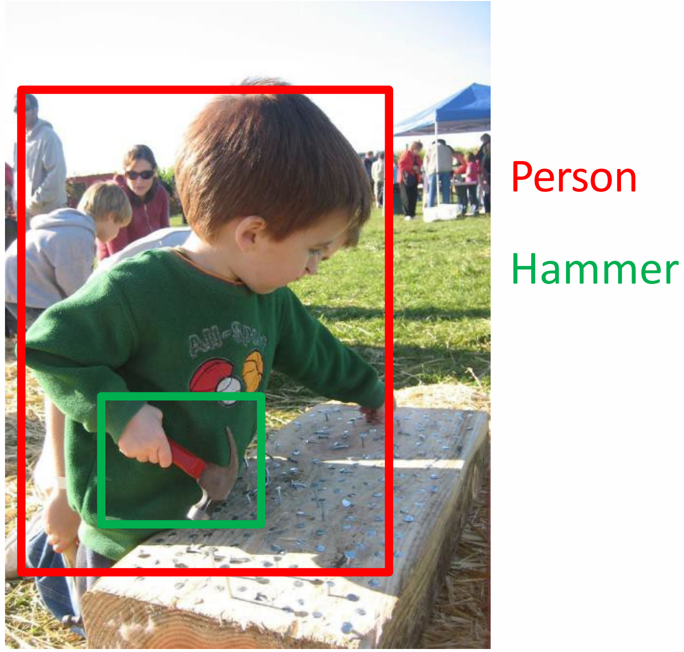
Image by Christina C. is licensed under CC BY-SA 4.0

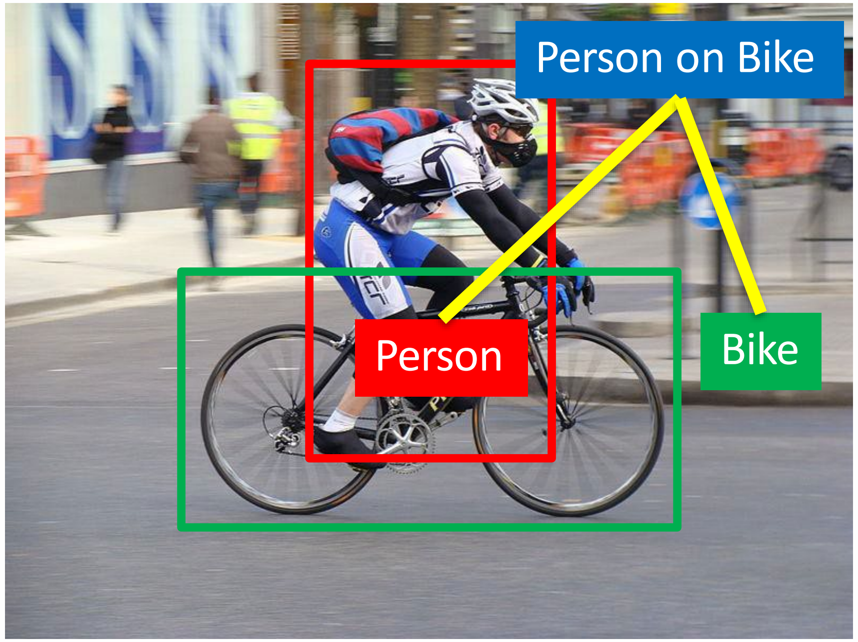
There are many visual recognition problems that are related to image classification, such as

object detection, image captioning



- Object detection
- Action classification
- Image captioning
- …



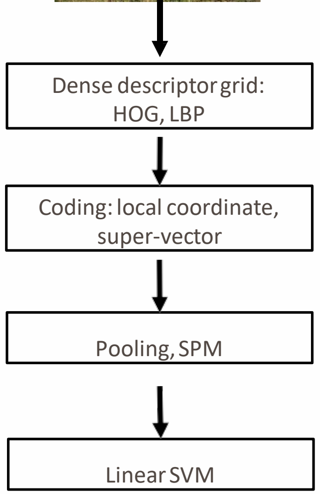


Convolutional Neural Networks (CNN) have become an important tool for object recognition

IMAGENET Large Scale Visual Recognition Challenge

Year 2010

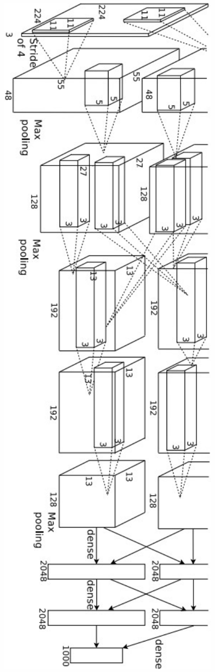
NEC-UIUC



[Lin CVPR 2011]

Year 2012

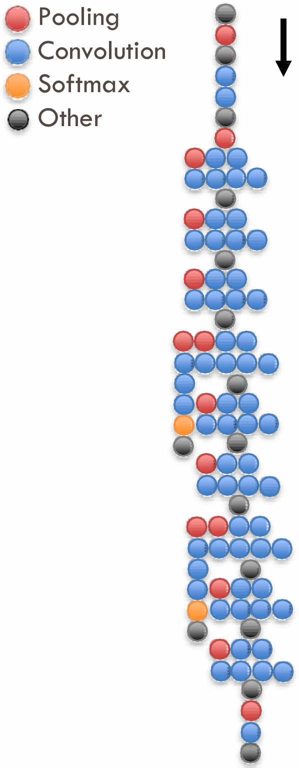
SuperVision



[Krizhevsky NIPS 2012]

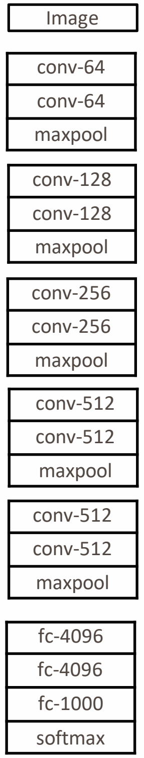
Year 2014

GoogLeNet



[Szegedy arxiv 2014]

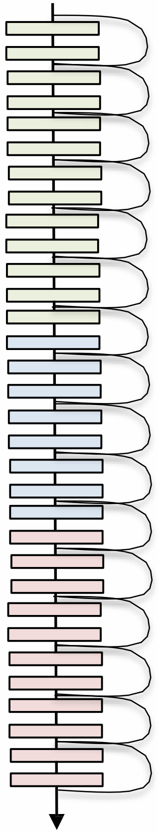
VGG



[Simonyan arxiv 2014]

Year 2015

MSRA



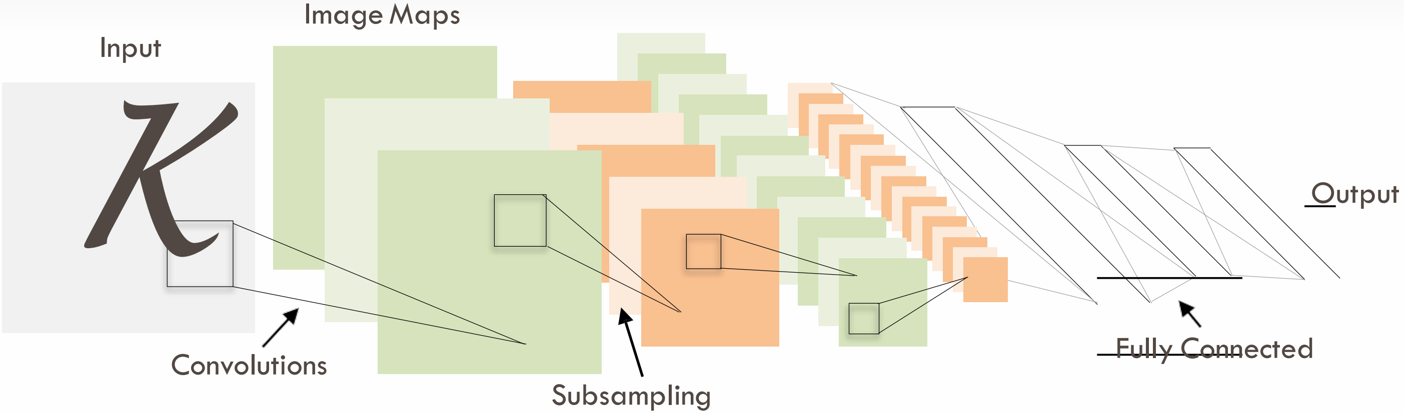
[He ICCV 2015]

Convolutional Neural Networks (CNN)

were not invented overnight

1998

LeCun et al.



# of transistors

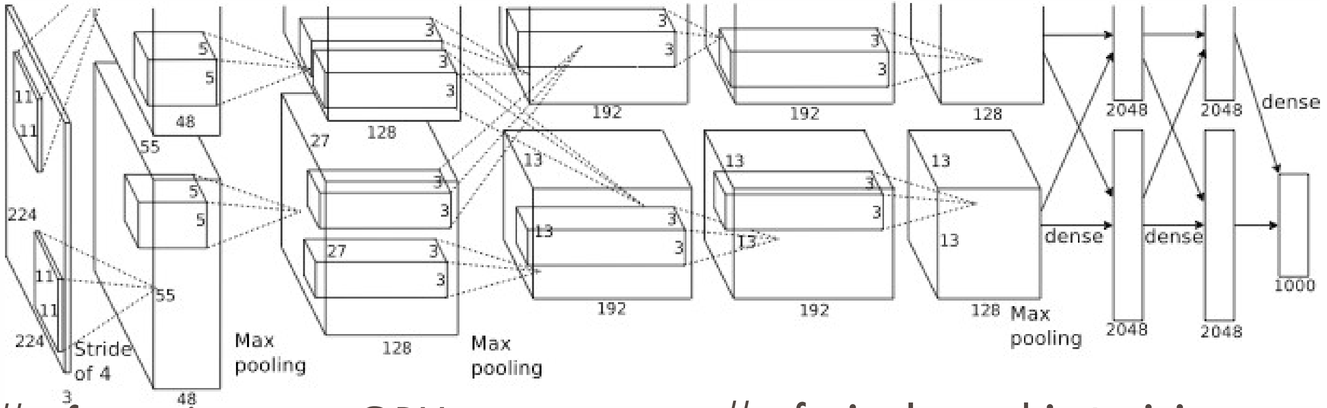
10^6

# of pixels used in training

10^7 NIST

2012

Krizhevsky et al.



# of transistors

10^9

GPUs

# of pixels used in training

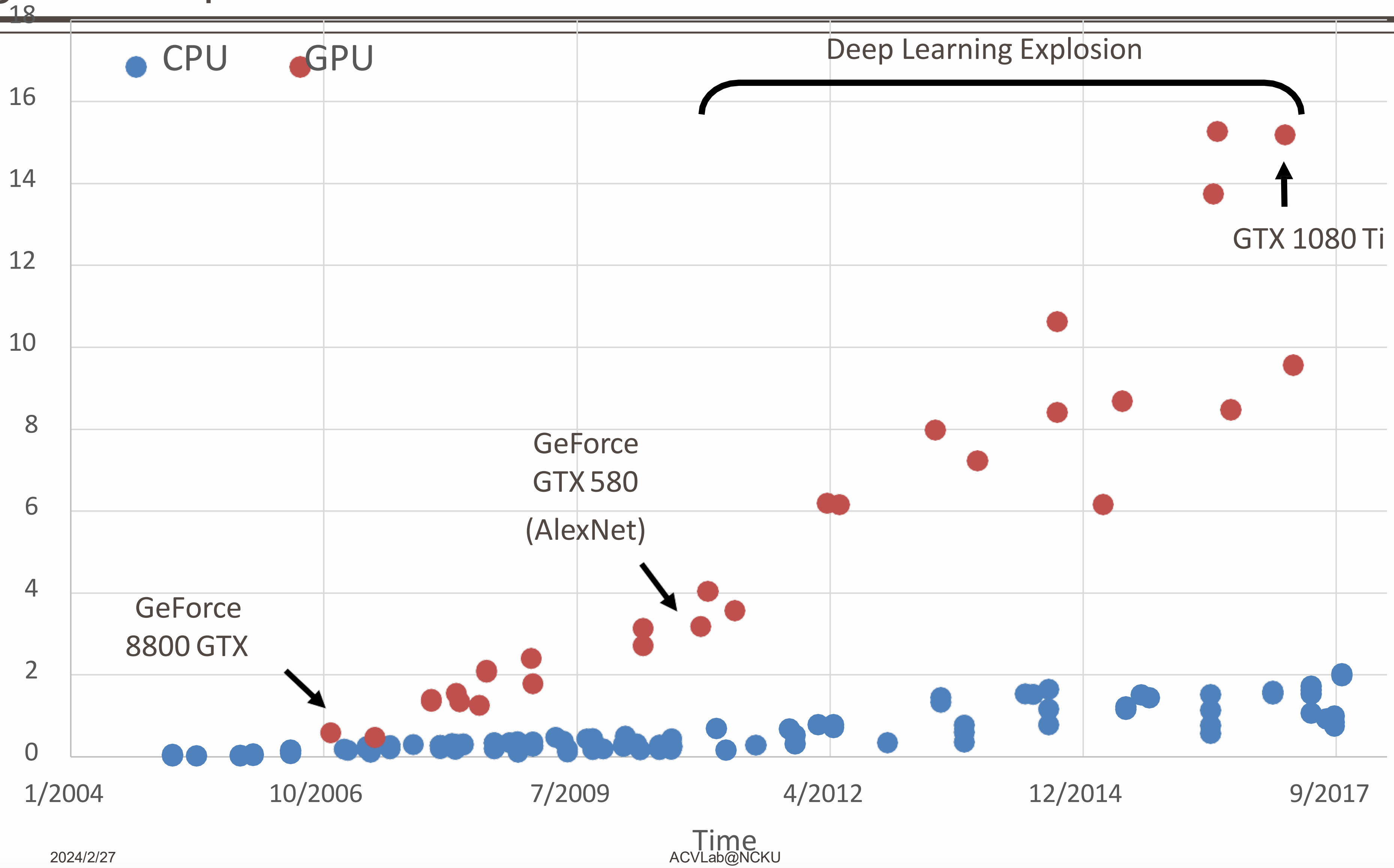
10^{14} IMAGENET

Figure copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

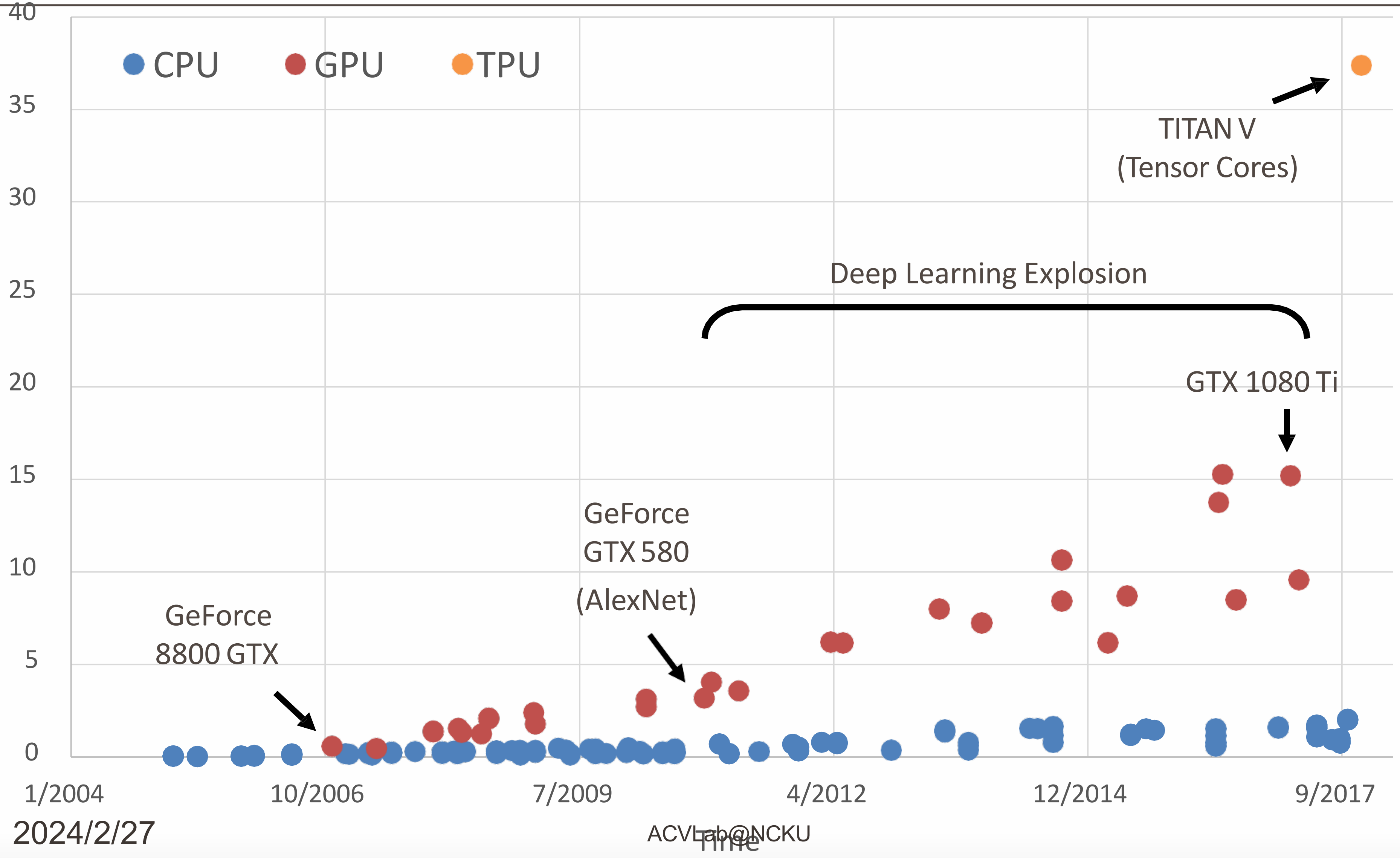
Ingredients for Deep Learning

GigaFLOPs per Dollar

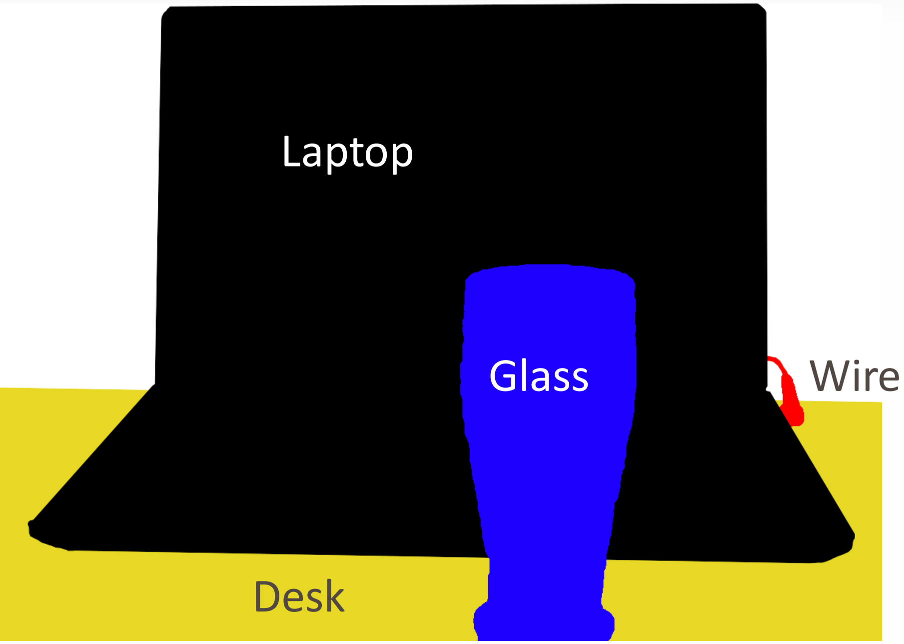
Floating-point operations per second

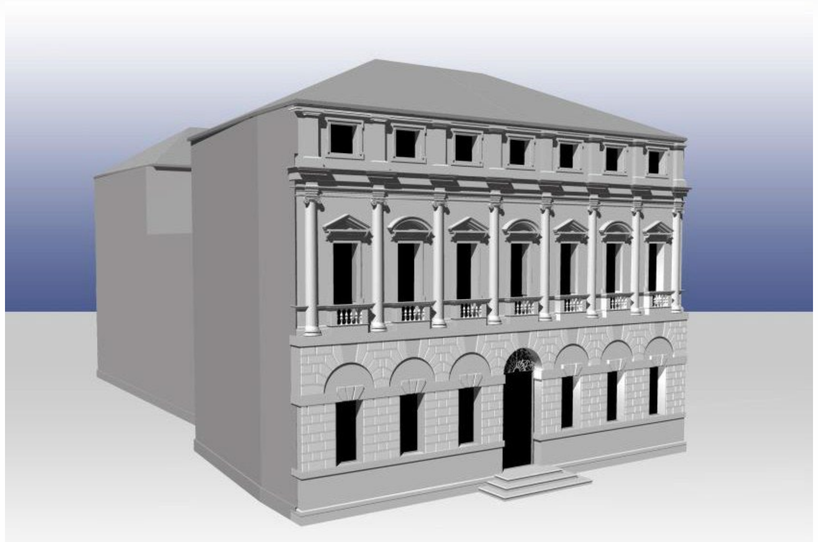


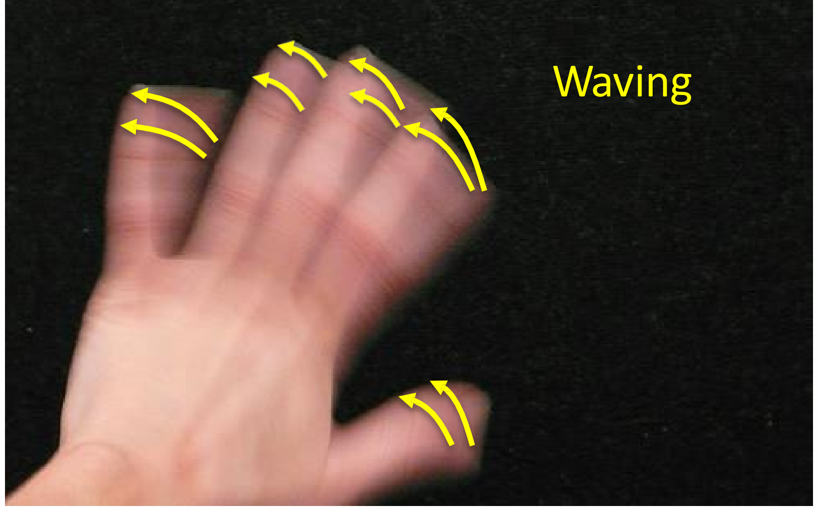
GigaFLOPs per Dollar



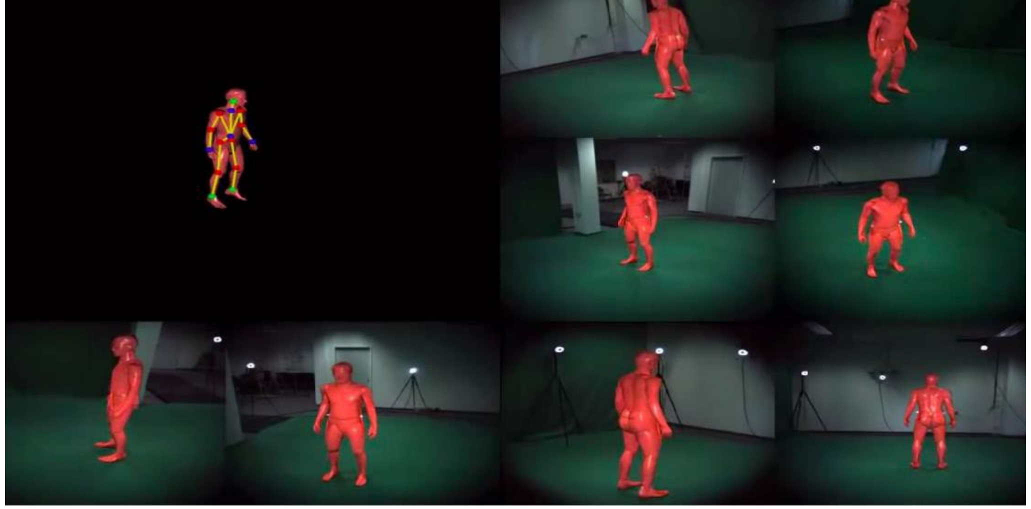
The quest for visual intelligence
goes far beyond object recognition…

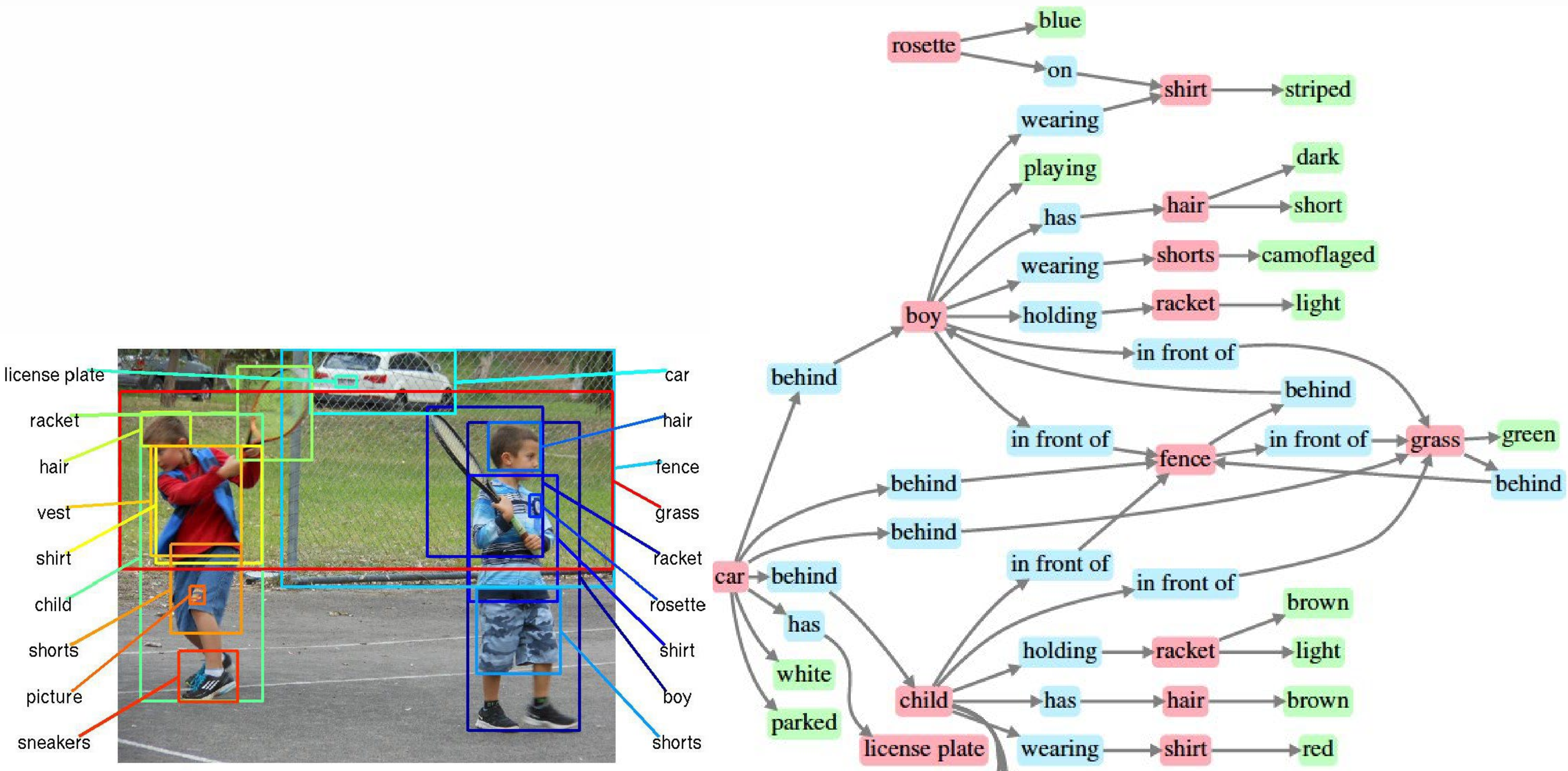

Wall
Laptop
Glass
Wire
Desk





Waving



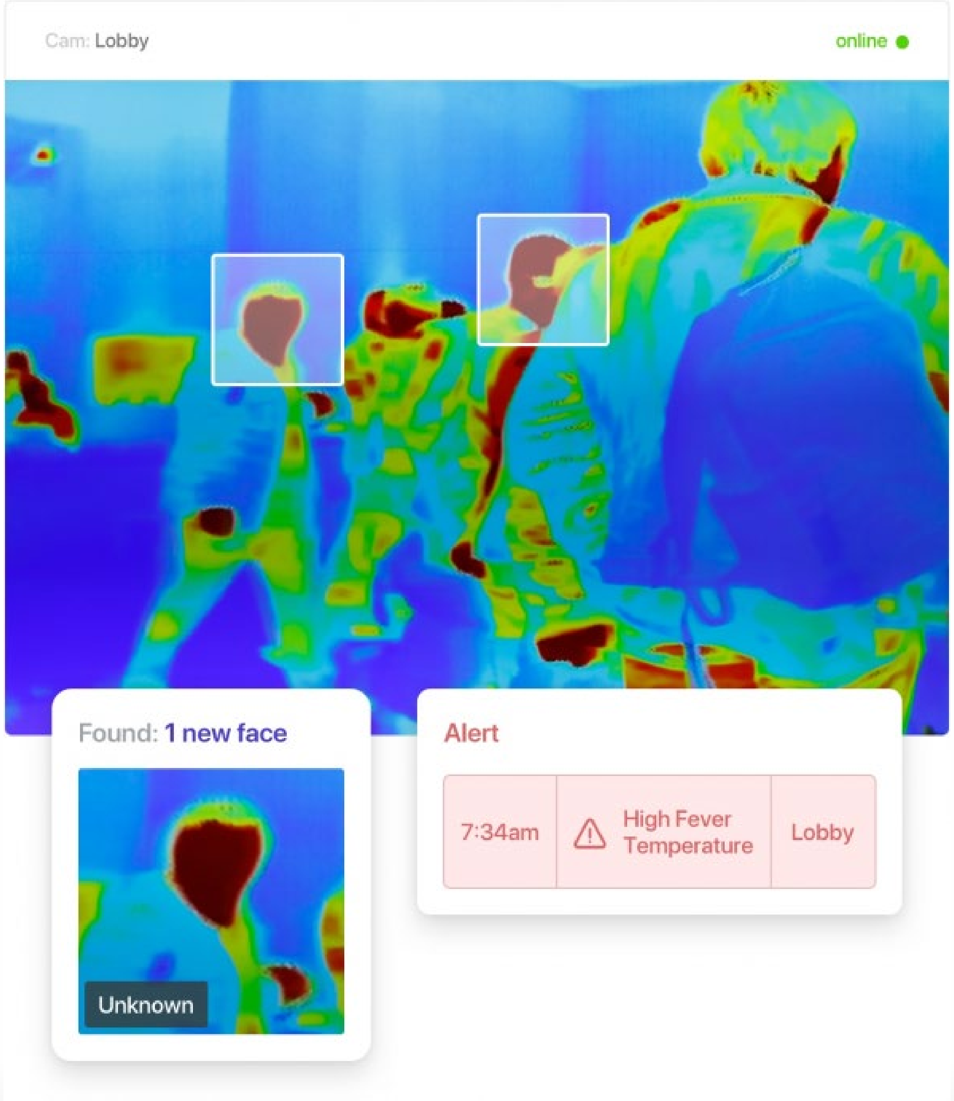


Johnson *et al.*, “Image Retrieval using Scene Graphs”, CVPR 2015

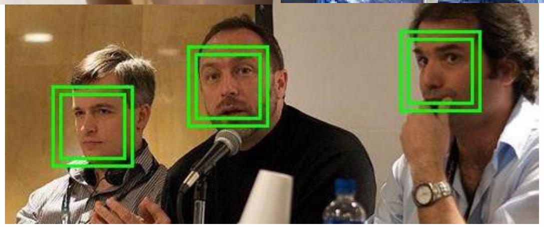
This image is copyright-free United States government work

Example credit: Andrej Karpathy

Face Recognition of NIR Images



Computer Vision Technology Can Better Our Lives



Outside border images, clockwise, starting from top left:
Image by Pop Culture Geek is licensed under CC BY 2.0; changes made
Image by the US Government is in the public domain
Image by the US Government is in the public domain
Image by Glogger is licensed under CC BY-SA 3.0; changes made
Image by Sylenius is licensed under CC BY 3.0; changes made
Image by US Government is in the public domain

Inside four images, clockwise, starting from top left:
Image is CC0 1.0 public domain
Image by Tucania is licensed under CC BY-SA 3.0; changes made
Image by Intuitive Surgical, Inc. is licensed under CC BY-SA 3.0; changes made
Image by Oyundari Zorigtbaatar is licensed under CC BY-SA 4.0

HOW?

IMAGE CLASSIFICATION PIPELINE

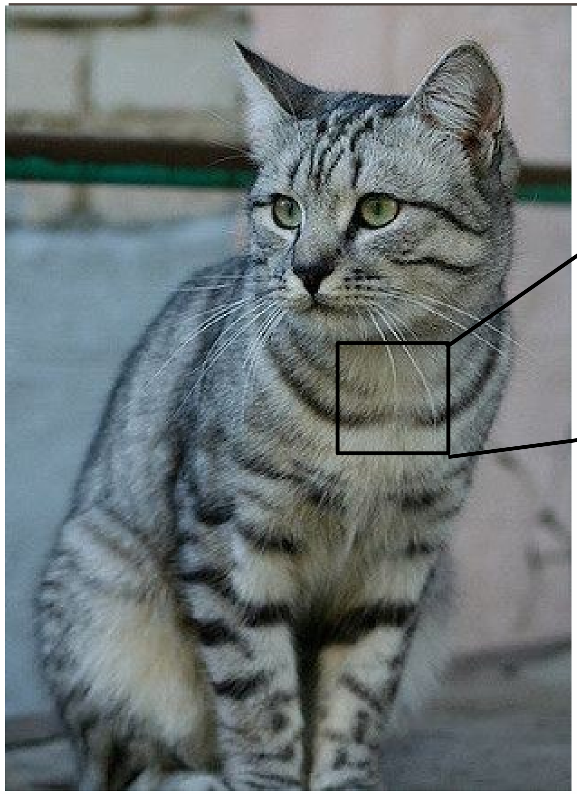
Image Classification: A core task in Computer Vision

This image by Nikita is licensed under CC-BY 2.0

(assume given a set of possible labels) {dog, cat, truck, plane, ...}

$\longrightarrow$ cat

The Problem: Semantic Gap



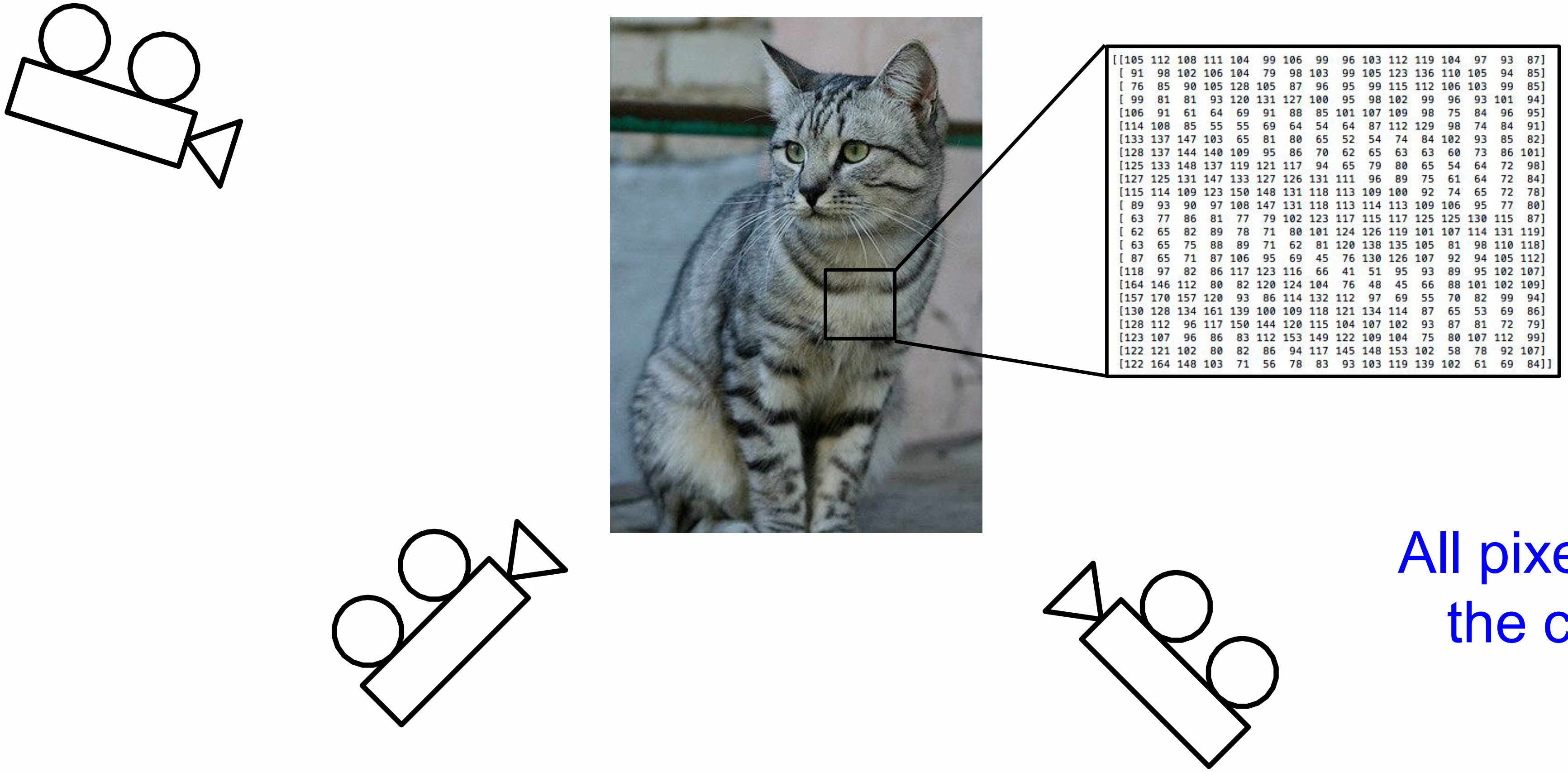

```
[[105 112 108 111 104  99 106  99  96 103 112 119 104  97  93  87]
 [ 91  98 102 106 104  79  98 103  99 105 123 136 110 105  94  85]
 [ 76  85  90 105 128 105  87  96  95  99 115 112 106 103  99  85]
 [ 99  81  81  93 120 131 127 100  95  98 102  99  96  93 101  94]
 [106  91  61  64  69  91  88  85 101 107 109  98  75  84  96  95]
 [114 108  85  55  55  69  64  54  64  87 112 129  98  74  84  91]
 [133 137 147 103  65  81  80  65  52  54  74  84 102  93  85  82]
 [128 137 144 140 109  95  86  70  62  65  63  63  60  73  86 101]
 [125 133 148 137 119 121 117  94  65  79  80  65  54  64  72  98]
 [127 125 131 147 133 127 126 131 111  96  89  75  61  64  72  84]
 [115 114 109 123 150 148 131 118 113 109 100  92  74  65  72  78]
 [ 89  93  90  97 108 147 131 118 113 114 113 109 106  95  77  80]
 [ 63  77  86  81  77  79 102 123 117 115 117 125 125 130 115  87]
 [ 62  65  82  89  78  71  80 101 124 126 119 101 107 114 131 119]
 [ 63  65  75  88  89  71  62  81 120 138 135 105  81  98 110 118]
 [ 87  65  71  87 106  95  69  45  76 130 126 107  92  94 105 112]
 [118  97  82  86 117 123 116  66  41  51  95  93  89  95 102 107]
 [164 146 112  80  82 120 124 104  76  48  45  66  88 101 102 109]
 [157 170 157 120  93  86 114 132 112  97  69  55  70  82  99  94]
 [130 128 134 161 139 100 109 118 121 134 114  87  65  53  69  86]
 [128 112  96 117 150 144 120 115 104 107 102  93  87  81  72  79]
 [123 107  96  86  83 112 153 149 122 109 104  75  80 107 112  99]
 [122 121 102  80  82  86  94 117 145 148 153 102  58  78  92 107]
 [122 164 148 103  71  56  78  83  93 103 119 139 102  61  69  84]]
```

What the computer sees

An image is a tensor of integers between [0, 255]:

e.g. 800 x 600 x 3
(3 channels RGB)

Challenges: Viewpoint variation



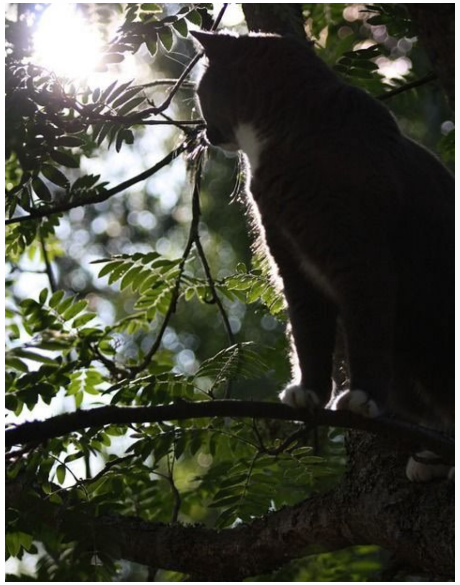
All pixels change when
the camera moves!

Challenges: Illumination


This image is CC0 1.0 public domain


This image is CC0 1.0 public domain


This image is CC0 1.0 public domain


This image is CC0 1.0 public domain

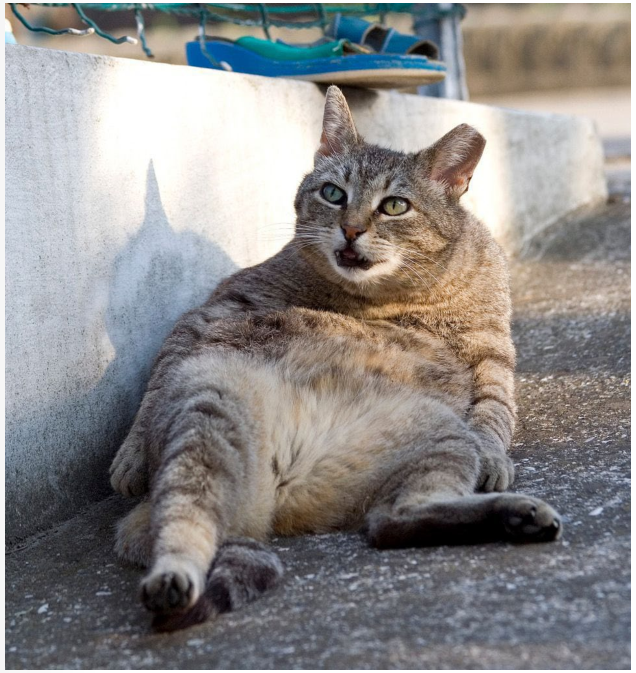
Challenges: Background Clutter

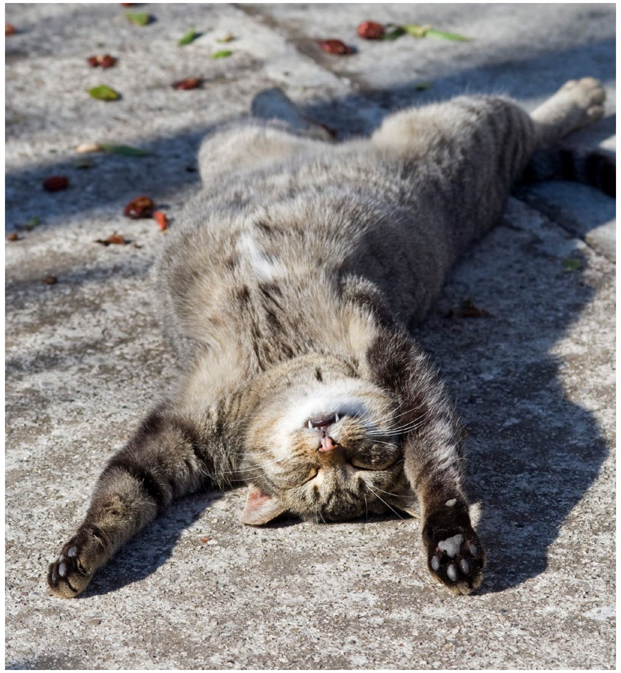
This image is CC0 1.0 public domain

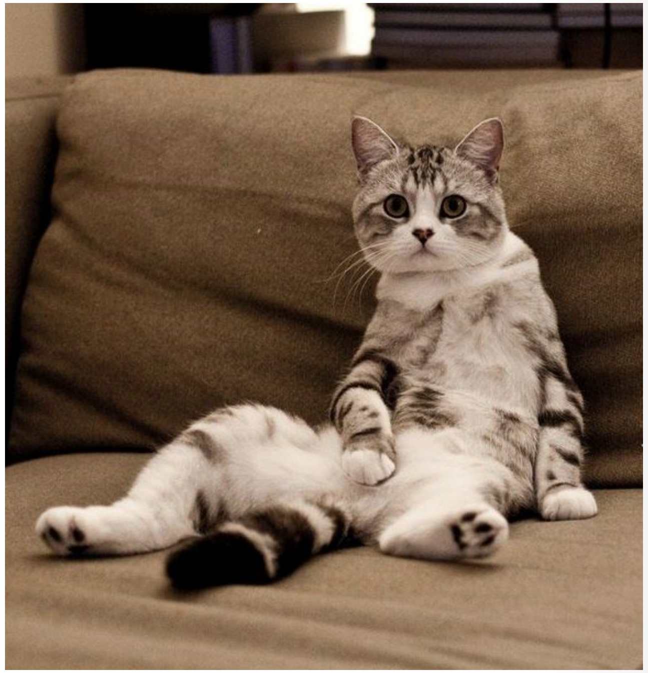
This image is CC0 1.0 public domain

Challenges: Occlusion

Challenges: Deformation







Challenges: Intraclass variation

Challenges: Context

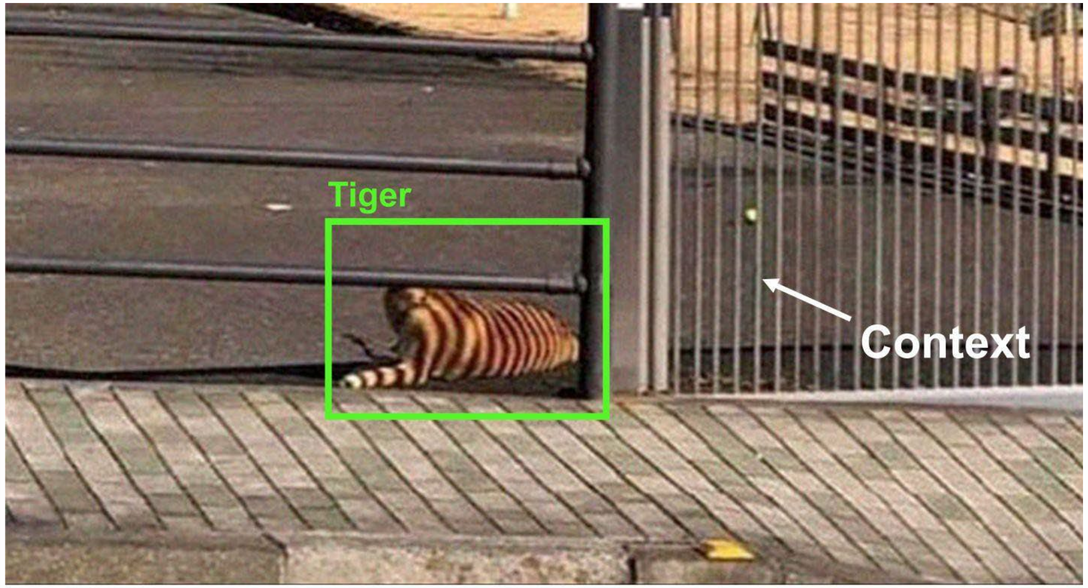


Image source:
https://www.linkedin.com/posts/ralph-aboujaoude-diaz-40838313_technology-artificialintelligence-computervision-activity-6912446088364875776-h-Iq?utm_source=linkedin_share&utm_medium=member_desktop_web

Modern computer vision algorithms

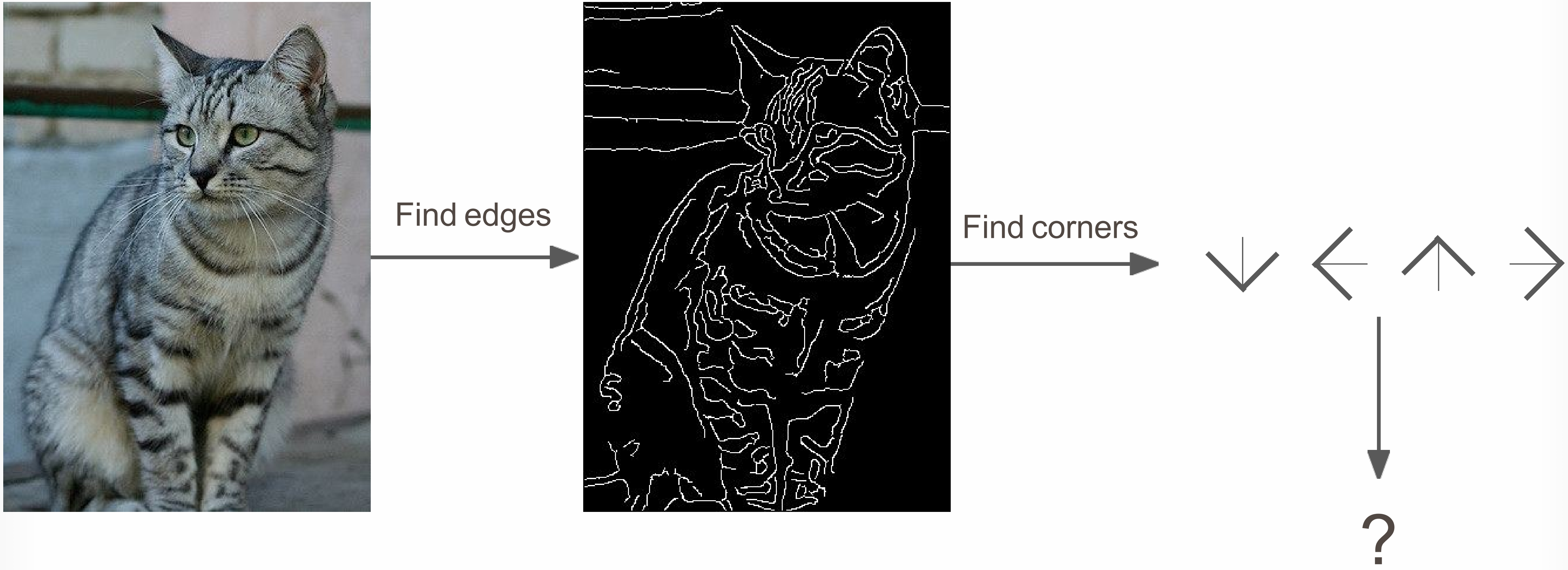
An image classifier

```
def classify_image(image):
  # Some magic here?
  return class_label
```

Unlike e.g. sorting a list of numbers,

no obvious way to hard-code the algorithm for recognizing a cat, or other classes.

Attempts have been made



John Canny, “A Computational Approach to Edge Detection”, IEEE TPAMI 1986

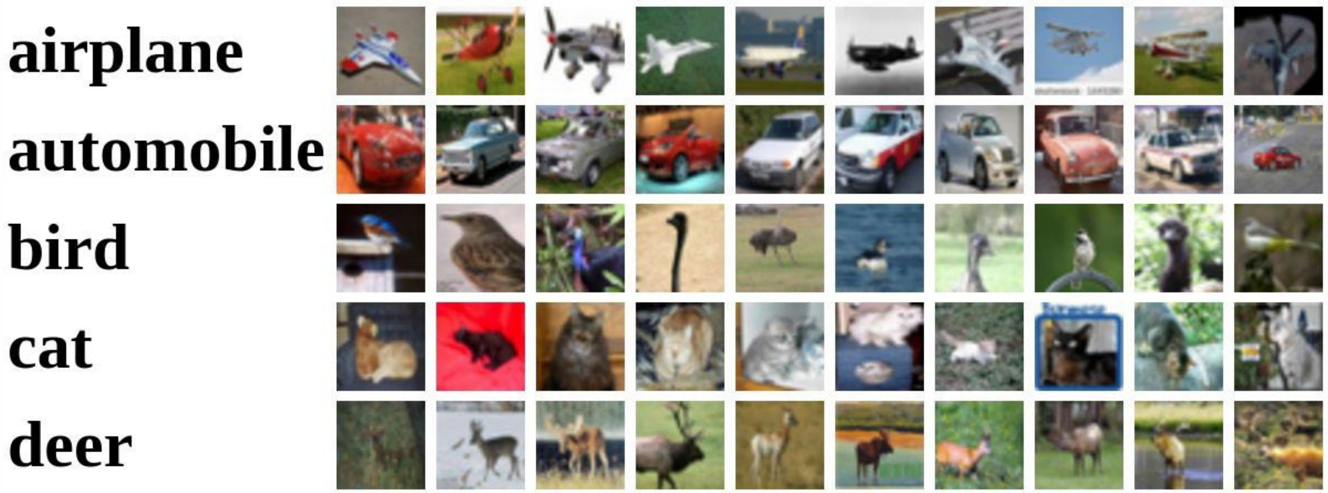
Machine Learning: Data-Driven Approach

1. Collect a dataset of images and labels
2. Use Machine Learning algorithms to train a classifier
3. Evaluate the classifier on new images

```python
def train(images, labels):
  # Machine learning!
  return model
```

```python
def predict(model, test_images):
  # Use model to predict labels
  return test_labels
```

- **Example training set**



NEAREST NEIGHBOR CLASSIFIER

First classifier: Nearest Neighbor

```
def train(images, labels):
  # Machine learning!
  return model
```

→ Memorize all data and labels

```
def predict(model, test_images):
  # Use model to predict labels
  return test_labels
```

→ Predict the label of the most similar training image

First classifier: Nearest Neighbor

deer 
bird 
plane 
cat 
car 

Training data with labels

?

query data

Distance Metric $|$  ,  $| \to \mathbb{R}$

Distance Metric to compare images

L1 distance: $d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$

$$
\left| \underbrace{\begin{array}{|c|c|c|c|} \hline 56 & 32 & 10 & 18 \\ \hline 90 & 23 & 128 & 133 \\ \hline 24 & 26 & 178 & 200 \\ \hline 2 & 0 & 255 & 220 \\ \hline \end{array}}_{\text{test image}} - \underbrace{\begin{array}{|c|c|c|c|} \hline 10 & 20 & 24 & 17 \\ \hline 8 & 10 & 89 & 100 \\ \hline 12 & 16 & 178 & 170 \\ \hline 4 & 32 & 233 & 112 \\ \hline \end{array}}_{\text{training image}} \right| = \underbrace{\begin{array}{|c|c|c|c|} \hline 46 & 12 & 14 & 1 \\ \hline 82 & 13 & 39 & 33 \\ \hline 12 & 10 & 0 & 30 \\ \hline 2 & 32 & 22 & 108 \\ \hline \end{array}}_{\text{pixel-wise absolute value differences}} \xrightarrow{\text{add}} 456
$$

Nearest Neighbor classifier

```python
import numpy as np

class NearestNeighbor:
  def __init__(self):
    pass

  def train(self, X, y):
    """ X is N x D where each row is an example. Y is 1-dimension of size N """
    # the nearest neighbor classifier simply remembers all the training data
    self.Xtr = X
    self.ytr = y

  def predict(self, X):
    """ X is N x D where each row is an example we wish to predict label for """
    num_test = X.shape[0]
    # lets make sure that the output type matches the input type
    Ypred = np.zeros(num_test, dtype = self.ytr.dtype)

    # loop over all test rows
    for i in xrange(num_test):
      # find the nearest training image to the i'th test image
      # using the L1 distance (sum of absolute value differences)
      distances = np.sum(np.abs(self.Xtr - X[i,:]), axis = 1)
      min_index = np.argmin(distances) # get the index with smallest distance
      Ypred[i] = self.ytr[min_index] # predict the label of the nearest example

    return Ypred
```

Nearest Neighbor classifier

```python
import numpy as np

class NearestNeighbor:
  def __init__(self):
    pass

  def train(self, X, y):
    """ X is N x D where each row is an example. Y is 1-dimension of size N """
    # the nearest neighbor classifier simply remembers all the training data
    self.Xtr = X
    self.ytr = y

  def predict(self, X):
    """ X is N x D where each row is an example we wish to predict label for """
    num_test = X.shape[0]
    # lets make sure that the output type matches the input type
    Ypred = np.zeros(num_test, dtype = self.ytr.dtype)

    # loop over all test rows
    for i in xrange(num_test):
      # find the nearest training image to the i'th test image
      # using the L1 distance (sum of absolute value differences)
      distances = np.sum(np.abs(self.Xtr - X[i,:]), axis = 1)
      min_index = np.argmin(distances) # get the index with smallest distance
      Ypred[i] = self.ytr[min_index] # predict the label of the nearest example

    return Ypred
```

Memorize training data

Nearest Neighbor classifier

```
import numpy as np

class NearestNeighbor:
  def __init__(self):
    pass

  def train(self, X, y):
    """ X is N x D where each row is an example. Y is 1-dimension of size N """
    # the nearest neighbor classifier simply remembers all the training data
    self.Xtr = X
    self.ytr = y

  def predict(self, X):
    """ X is N x D where each row is an example we wish to predict label for """
    num_test = X.shape[0]
    # lets make sure that the output type matches the input type
    Ypred = np.zeros(num_test, dtype = self.ytr.dtype)

    # loop over all test rows
    for i in xrange(num_test):
      # find the nearest training image to the i'th test image
      # using the L1 distance (sum of absolute value differences)
      distances = np.sum(np.abs(self.Xtr - X[i,:]), axis = 1)
      min_index = np.argmin(distances) # get the index with smallest distance
      Ypred[i] = self.ytr[min_index] # predict the label of the nearest example

    return Ypred
```

For each test image:
  Find closest train image
  Predict label of nearest image

```python
import numpy as np

class NearestNeighbor:
  def __init__(self):
    pass

  def train(self, X, y):
    """ X is N x D where each row is an example. Y is 1-dimension of size N """
    # the nearest neighbor classifier simply remembers all the training data
    self.Xtr = X
    self.ytr = y

  def predict(self, X):
    """ X is N x D where each row is an example we wish to predict label for """
    num_test = X.shape[0]
    # lets make sure that the output type matches the input type
    Ypred = np.zeros(num_test, dtype = self.ytr.dtype)

    # loop over all test rows
    for i in xrange(num_test):
      # find the nearest training image to the i'th test image
      # using the L1 distance (sum of absolute value differences)
      distances = np.sum(np.abs(self.Xtr - X[i,:]), axis = 1)
      min_index = np.argmin(distances) # get the index with smallest distance
      Ypred[i] = self.ytr[min_index] # predict the label of the nearest example

    return Ypred
```

Nearest Neighbor classifier

Q: With N examples, how fast are training and prediction?

Ans: Train O(1),
predict O(N)

This is bad: we want classifiers that are **fast** at prediction; **slow** for training is ok

```python
import numpy as np

class NearestNeighbor:
  def __init__(self):
    pass

  def train(self, X, y):
    """ X is N x D where each row is an example. Y is 1-dimension of size N """
    # the nearest neighbor classifier simply remembers all the training data
    self.Xtr = X
    self.ytr = y

  def predict(self, X):
    """ X is N x D where each row is an example we wish to predict label for """
    num_test = X.shape[0]
    # lets make sure that the output type matches the input type
    Ypred = np.zeros(num_test, dtype = self.ytr.dtype)

    # loop over all test rows
    for i in xrange(num_test):
      # find the nearest training image to the i'th test image
      # using the L1 distance (sum of absolute value differences)
      distances = np.sum(np.abs(self.Xtr - X[i,:]), axis = 1)
      min_index = np.argmin(distances) # get the index with smallest distance
      Ypred[i] = self.ytr[min_index] # predict the label of the nearest example

    return Ypred
```

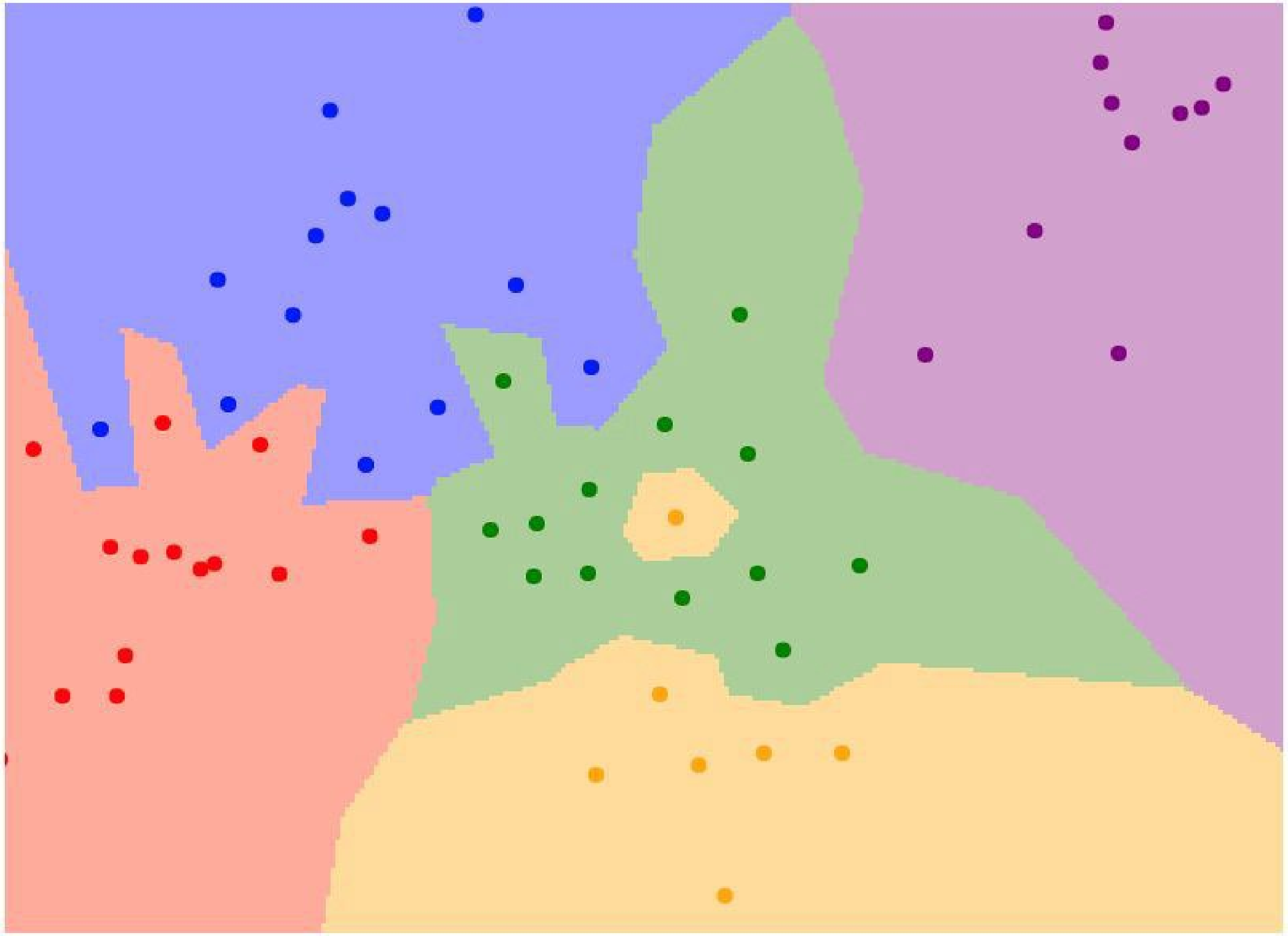
Nearest Neighbor classifier

Many methods exist for fast / approximate nearest neighbor

A good implementation:
https://github.com/facebookresearch/faiss

Johnson et al, "Billion-scale similarity search with GPUs", arXiv 2017

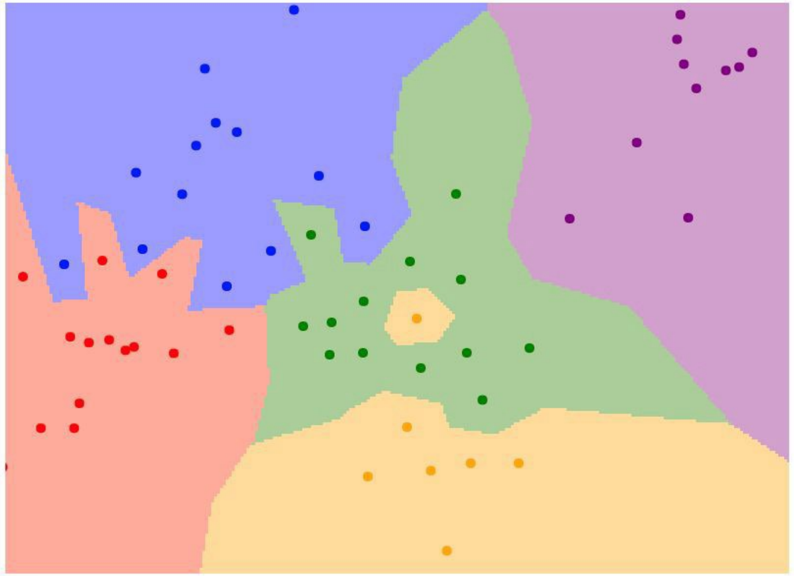
What does this look like?



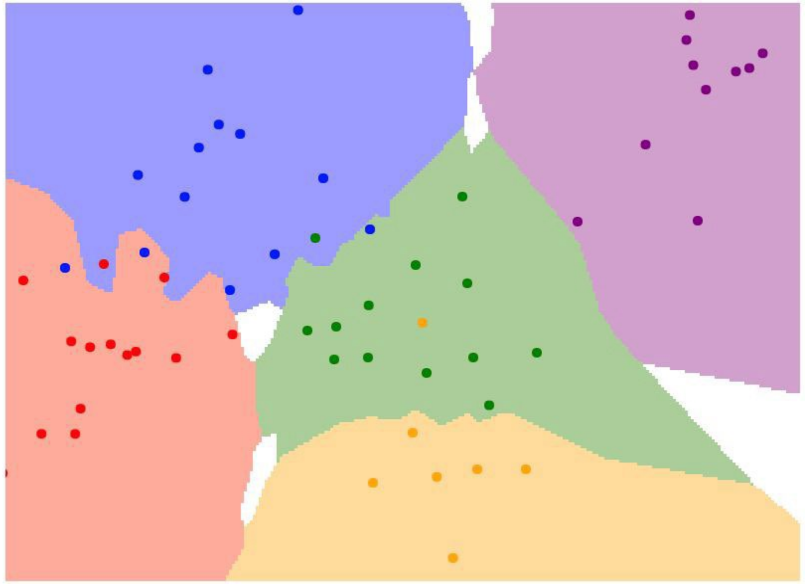
1-nearest neighbor

K-Nearest Neighbors

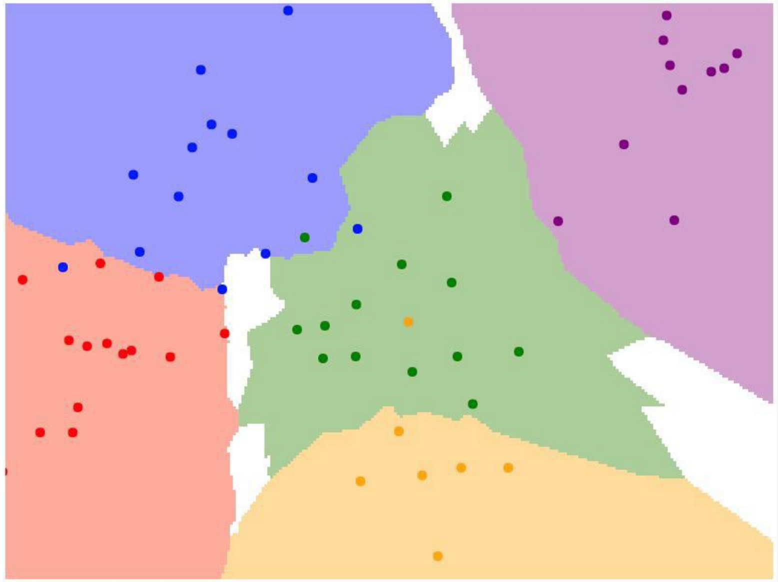
Instead of copying label from nearest neighbor,
take **majority vote** from K closest points



K = 1



K = 3

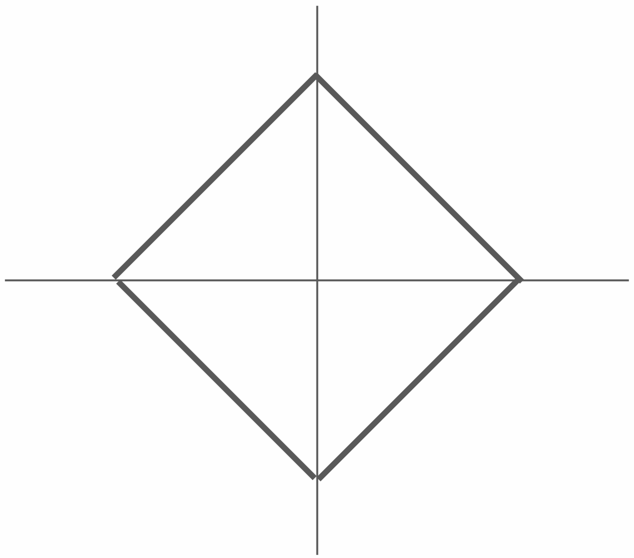


K = 5

K-Nearest Neighbors: Distance Metric

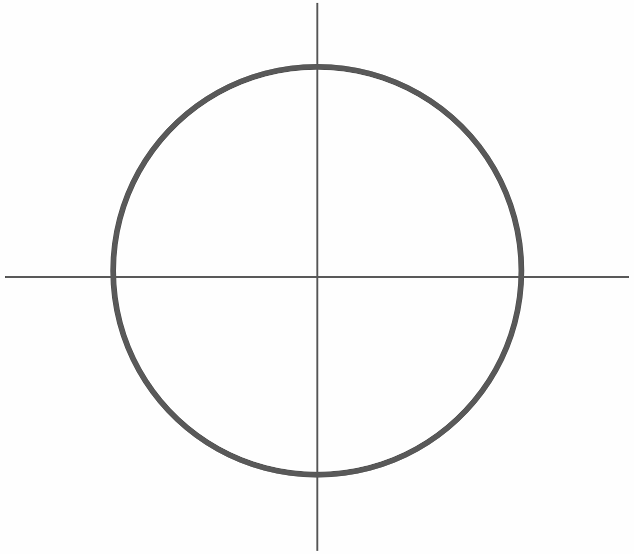
L1 (Manhattan) distance

$$d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$$



L2 (Euclidean) distance

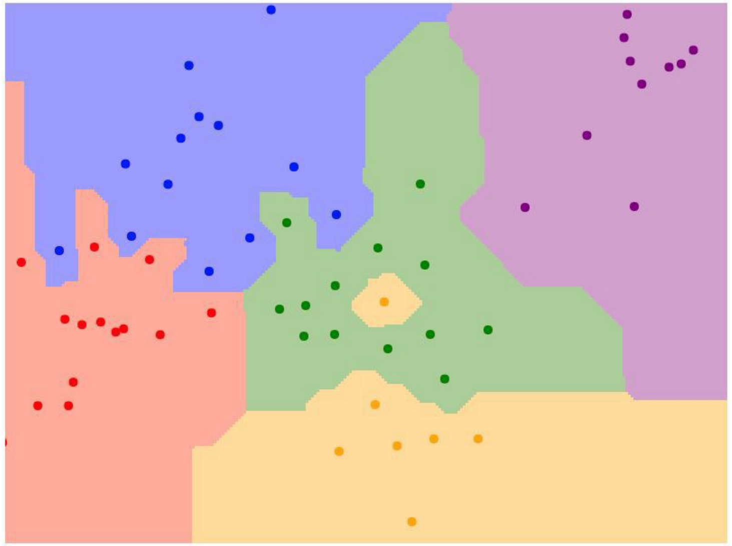
$$d_2(I_1, I_2) = \sqrt{\sum_p (I_1^p - I_2^p)^2}$$



K-Nearest Neighbors: Distance Metric

L1 (Manhattan) distance

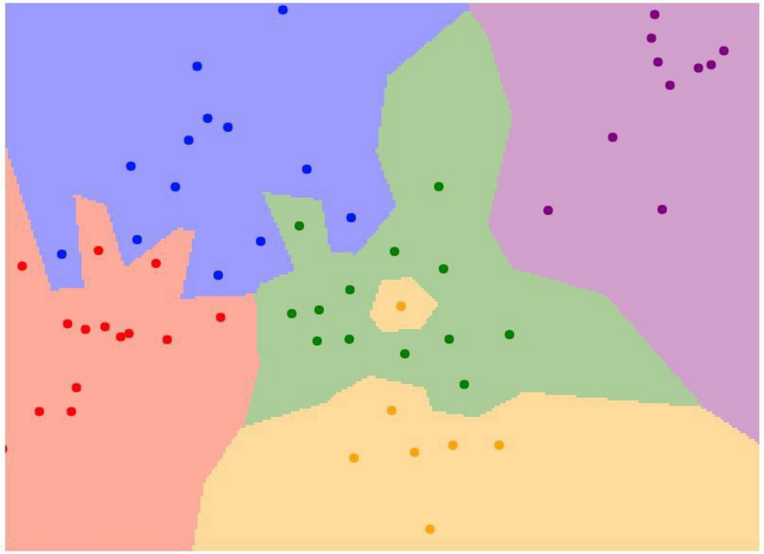
$$d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$$



K = 1

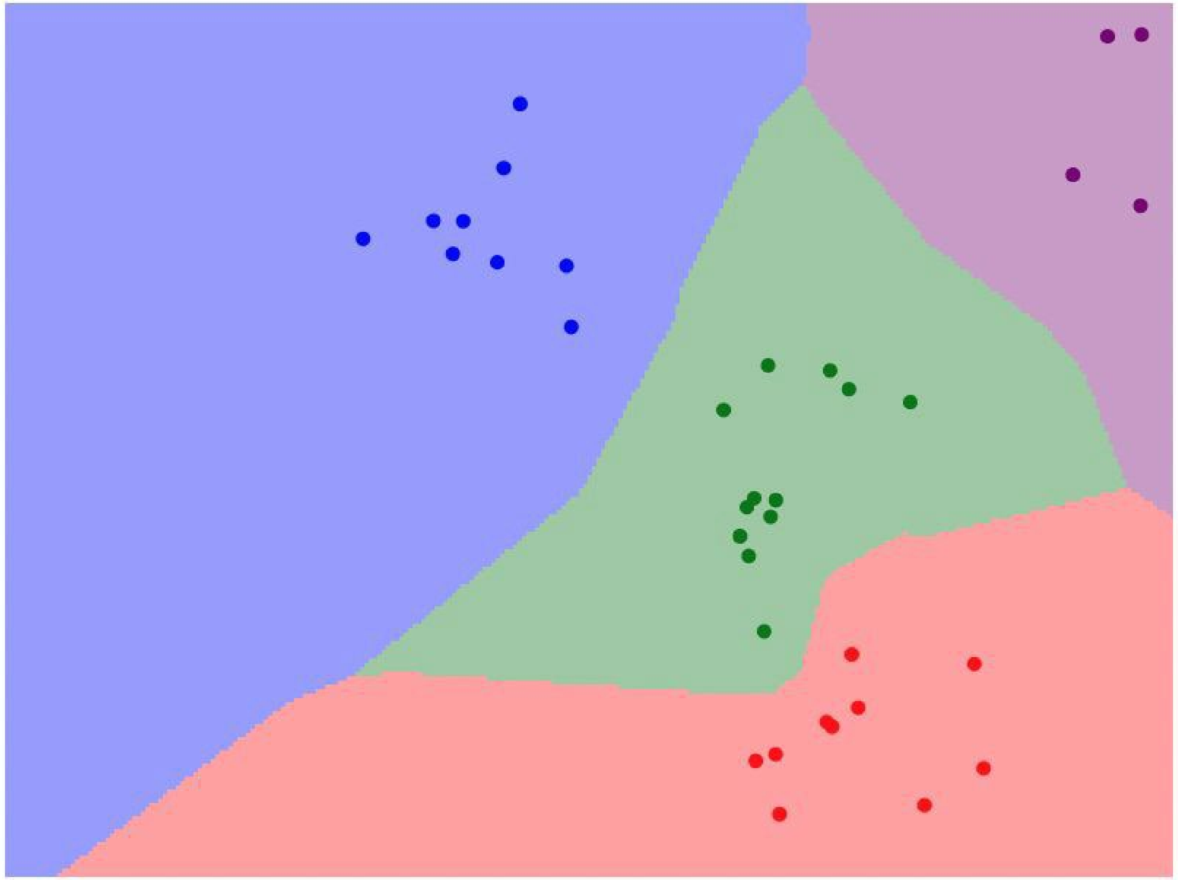
L2 (Euclidean) distance

$$d_2(I_1, I_2) = \sqrt{\sum_p \left(I_1^p - I_2^p\right)^2}$$



K = 1

K-Nearest Neighbors: try it yourself!



http://vision.stanford.edu/teaching/cs231n-demos/knn/

Hyperparameters

- What is the best value of **k** to use? What is the best distance to use?

- These are **hyperparameters**: choices about the algorithms themselves.

- Very problem/dataset-dependent.
- Must try them all out and see what works best.

Setting Hyperparameters

Idea #1: Choose hyperparameters that work best on the **training data**

BAD: K = 1 always works perfectly on training data

train

Idea #2: choose hyperparameters that work best on **test** data

BAD: No idea how algorithm will perform on new data

train	test

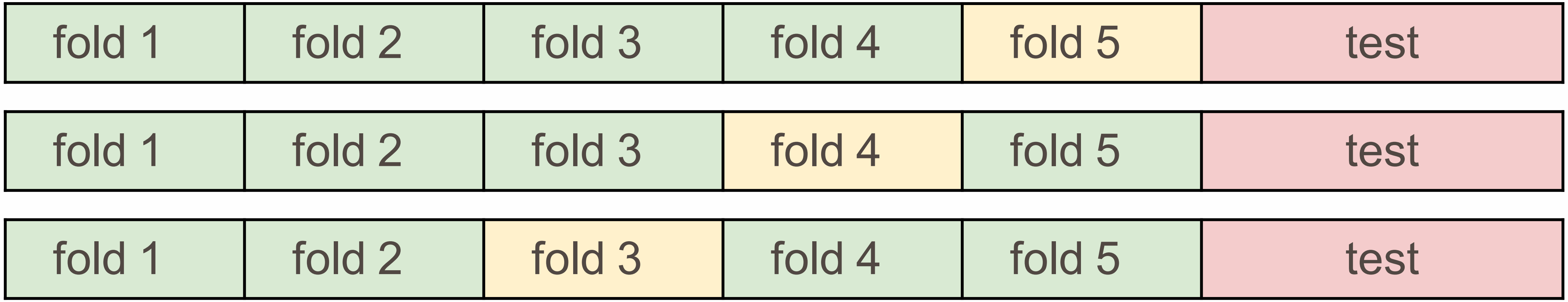
Idea #3: Split data into **train**, **val**; choose hyperparameters on val and evaluate on test

Better!

train	validation	test

Setting Hyperparameters

Idea #4: **Cross-Validation**: Split data into **folds**,
try each fold as validation and average the results



Useful for small datasets, but not used too frequently in deep learning

Example Dataset: CIFAR10

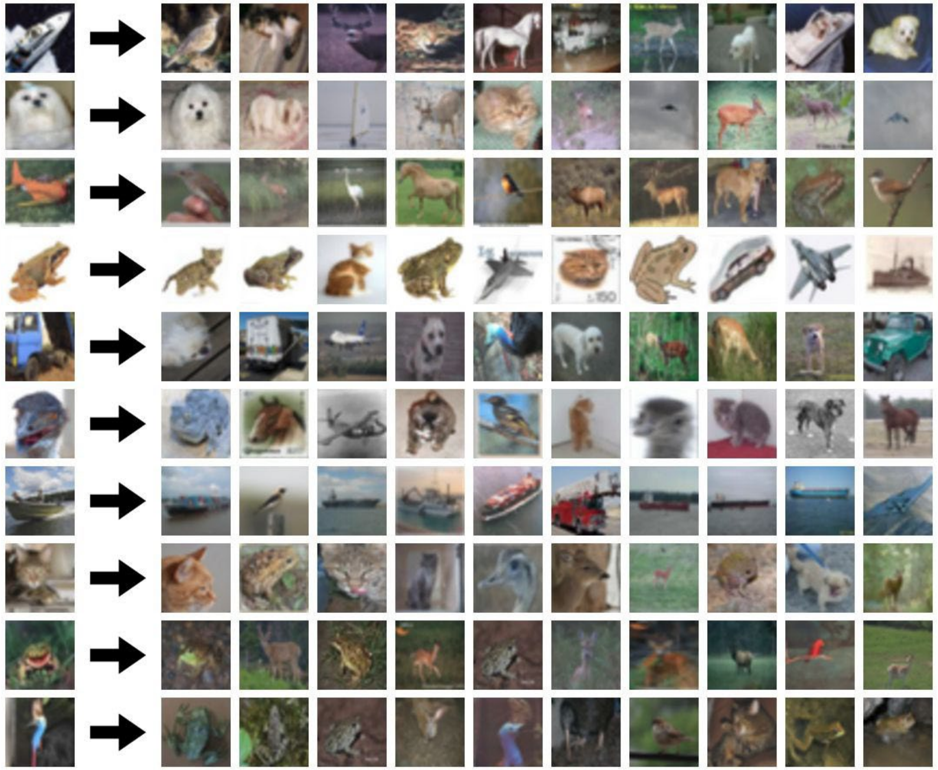
10 classes
50,000 training images
10,000 testing images

Alex Krizhevsky, "Learning Multiple Layers of Features from Tiny Images", Technical Report, 2009.

Example Dataset: CIFAR10

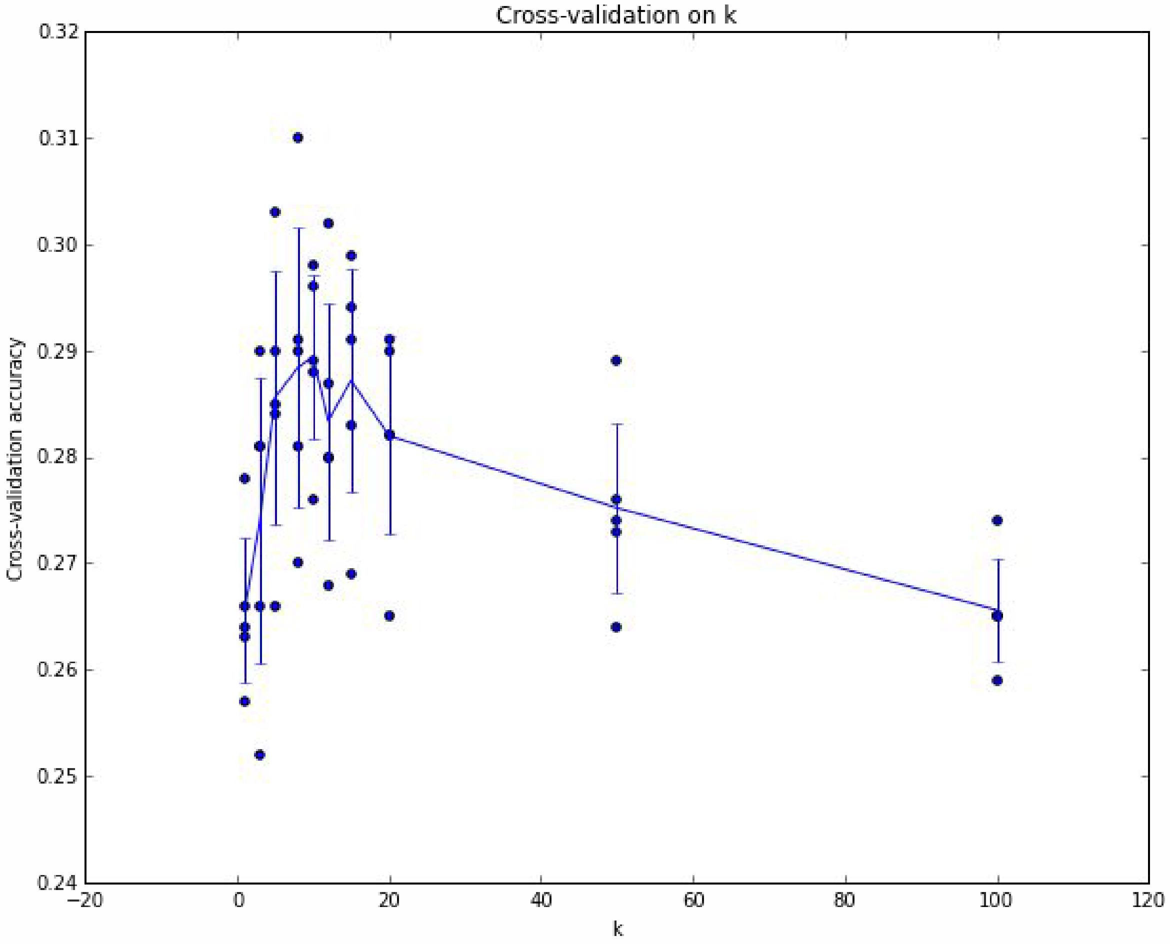
10 classes
50,000 training images
10,000 testing images

Test images and nearest neighbors



Alex Krizhevsky, “Learning Multiple Layers of Features from Tiny Images”, Technical Report, 2009.

Setting Hyperparameters



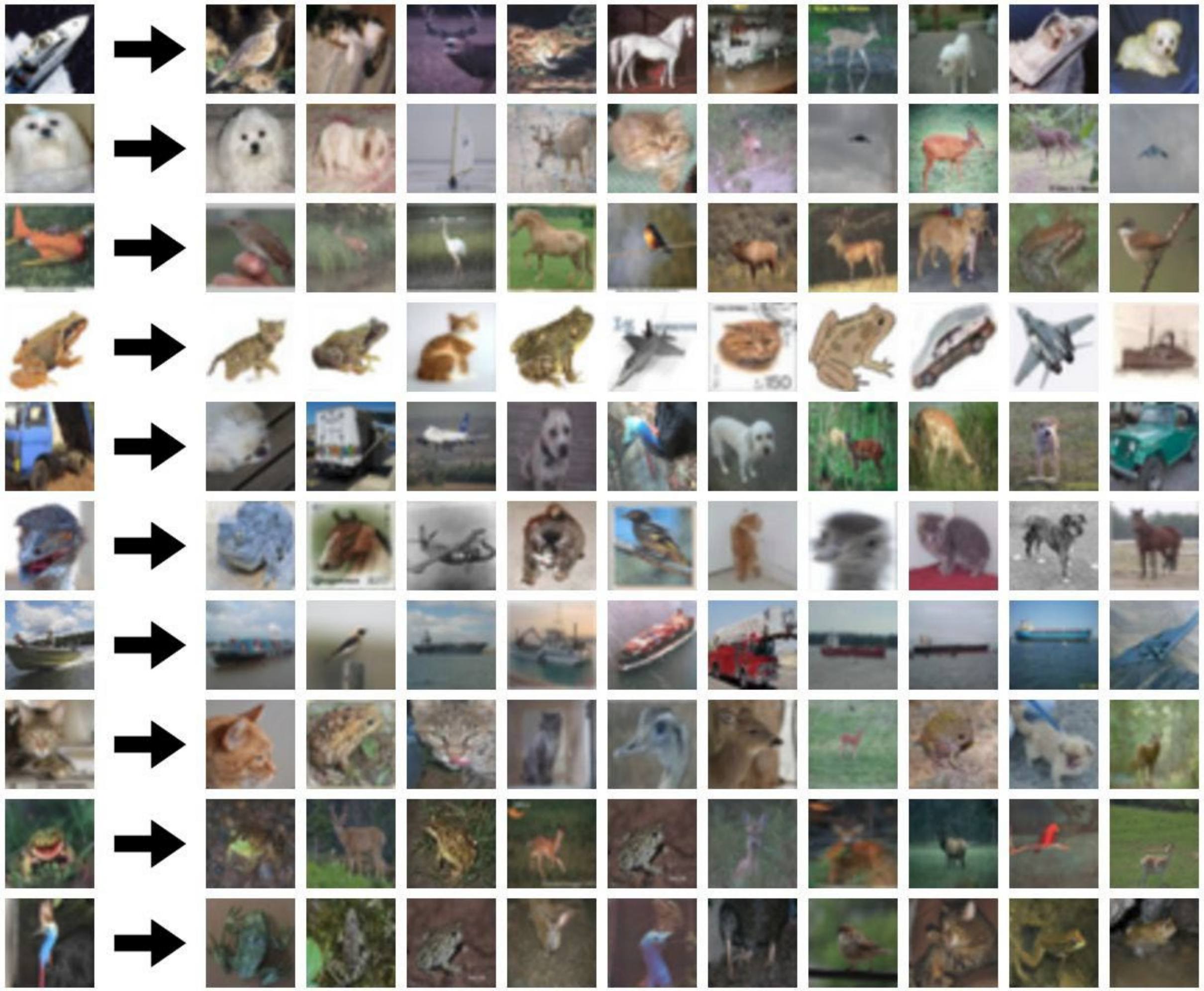
Example of
5-fold cross-validation for the value of k.

Each point: single outcome.

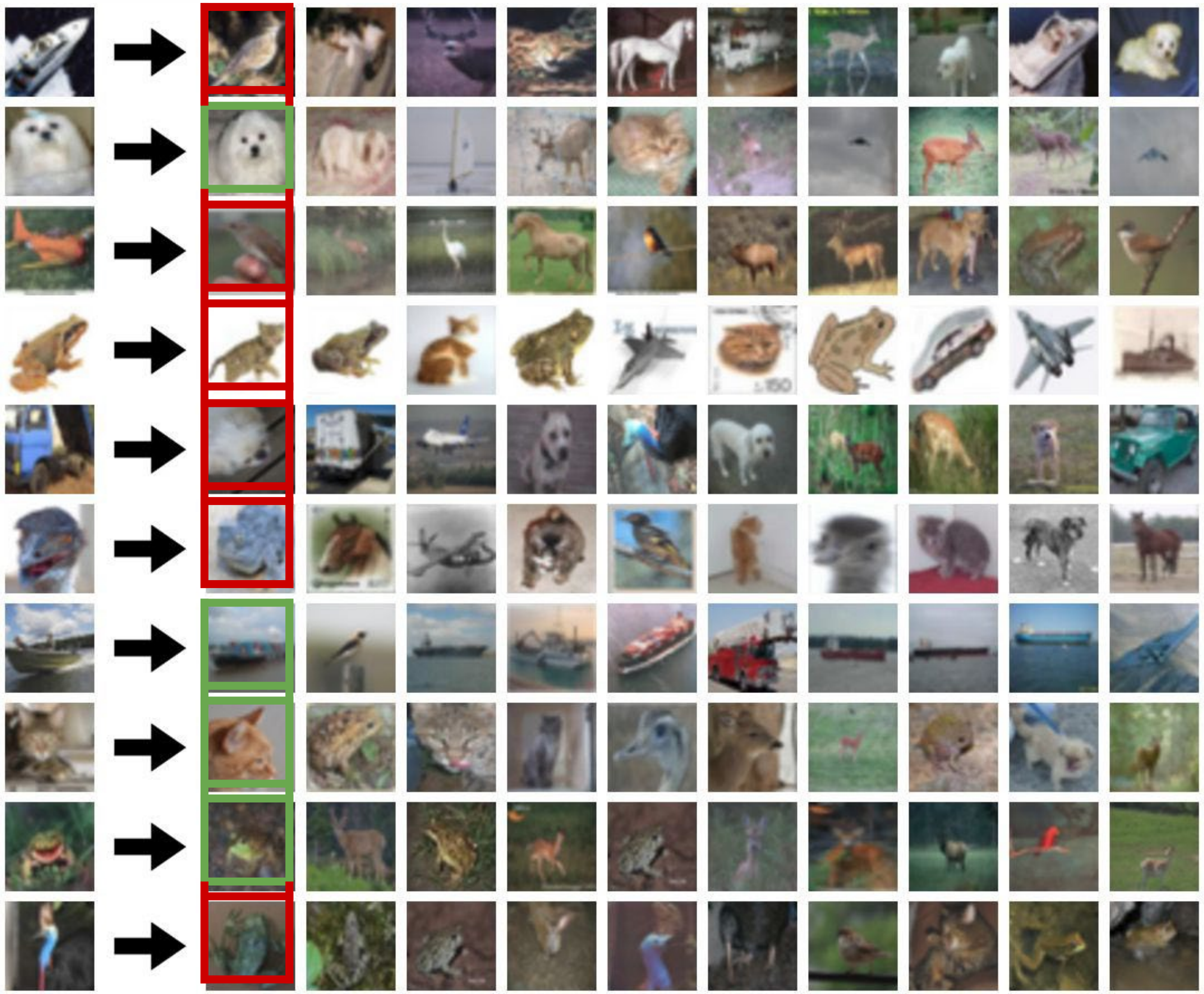
The line goes through the mean, bars indicated standard deviation

(Seems that k ~= 7 works best for this data)

What does this look like?



What does this look like?



k-Nearest Neighbor with pixel distance never used.

- Distance metrics on pixels are not informative

Original image is CC0 public domain

Original

Occluded

Shifted (1 pixel)

Tinted

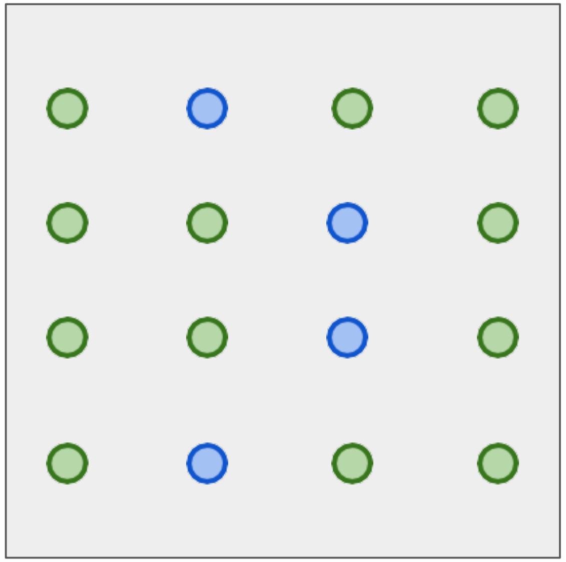
(All three images on the right have the same pixel distances to the one on the left)

k-Nearest Neighbor with pixel distance never used.

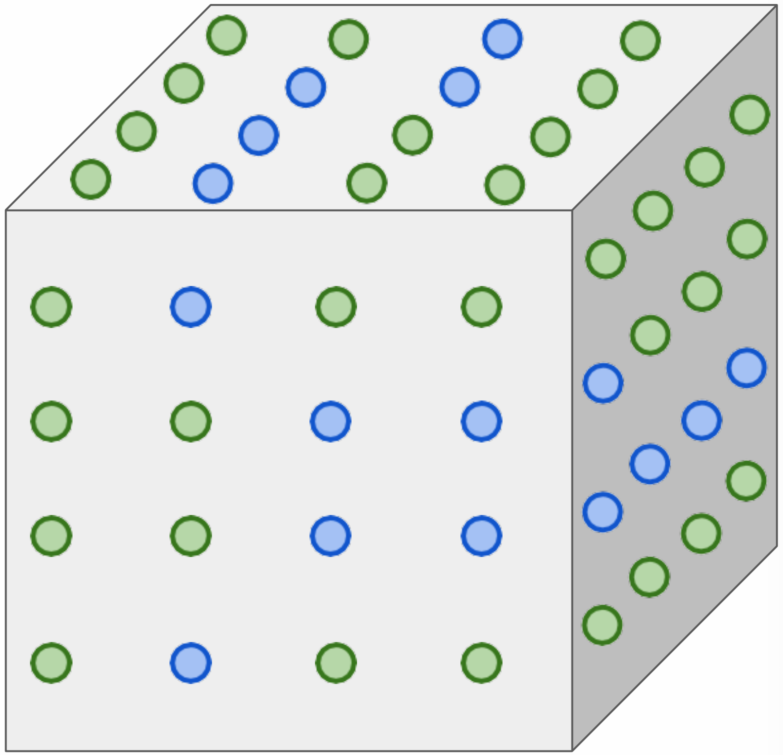
- Curse of dimensionality

Dimensions = 1
Points = 4

Dimensions = 2
Points = 4^2



Dimensions = 3
Points = 4^3



K-Nearest Neighbors: Summary

- In **image classification** we start with a **training set** of images and labels, and must predict labels on the **test set**

- The **K-Nearest Neighbors** classifier predicts labels based on the K nearest training examples

- Distance metric and K are **hyperparameters** Choose

hyperparameters using the **validation set** Only run on

the test set once at the very end!

LINEAR CLASSIfiER

Parametric Approach

3072x1

$$f(x,W) = \boxed{W}\boxed{x} + \boxed{b}$$

10x1

f(x,W) = Wx

10x1 10x3072

Image

Array of **32x32x3** numbers
(3072 numbers total)

$\longrightarrow$ f(**x**,**W**) $\longrightarrow$ **10** numbers giving class scores

$\uparrow$

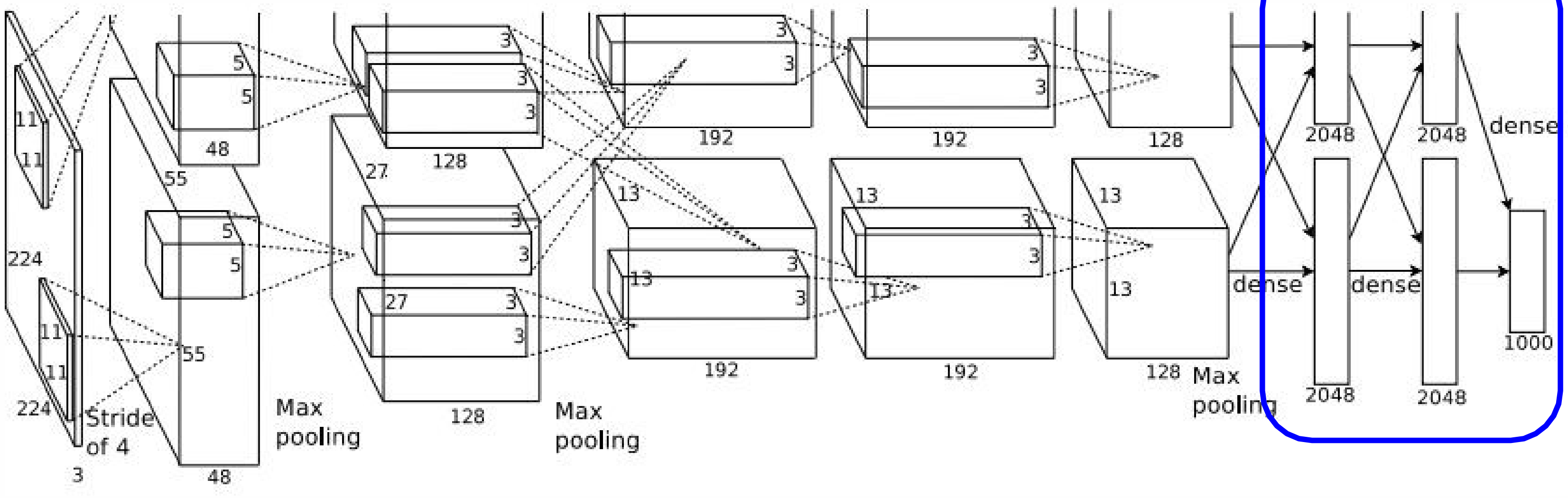
W

parameters
or weights

Neural Network

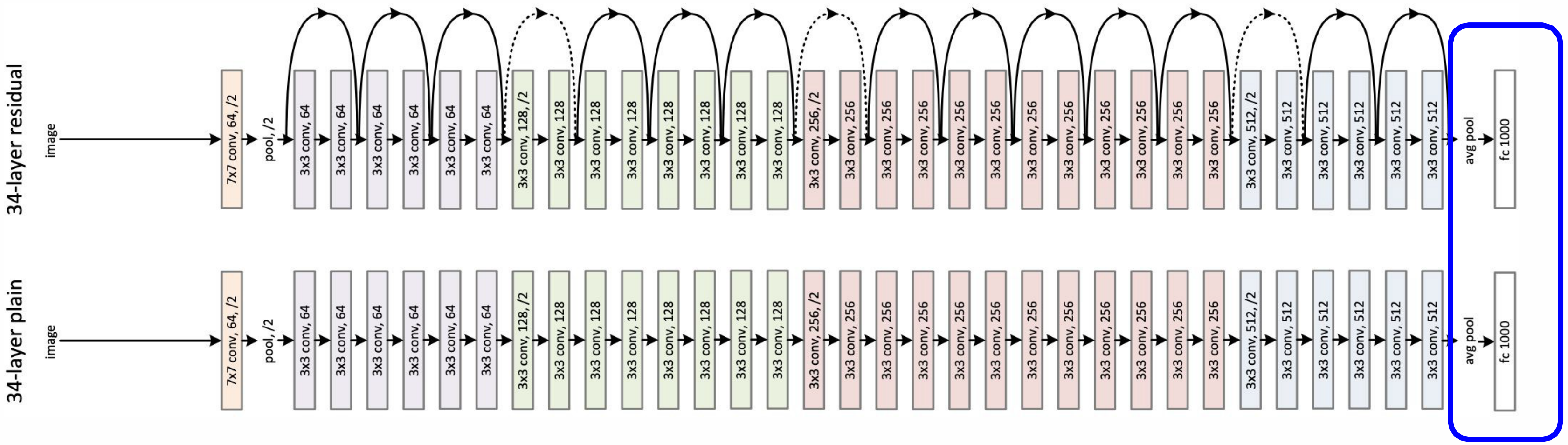
Linear
classifiers

This image is CC0 1.0 public domain



[Krizhevsky et al. 2012]

Linear layers



[He et al. 2015]

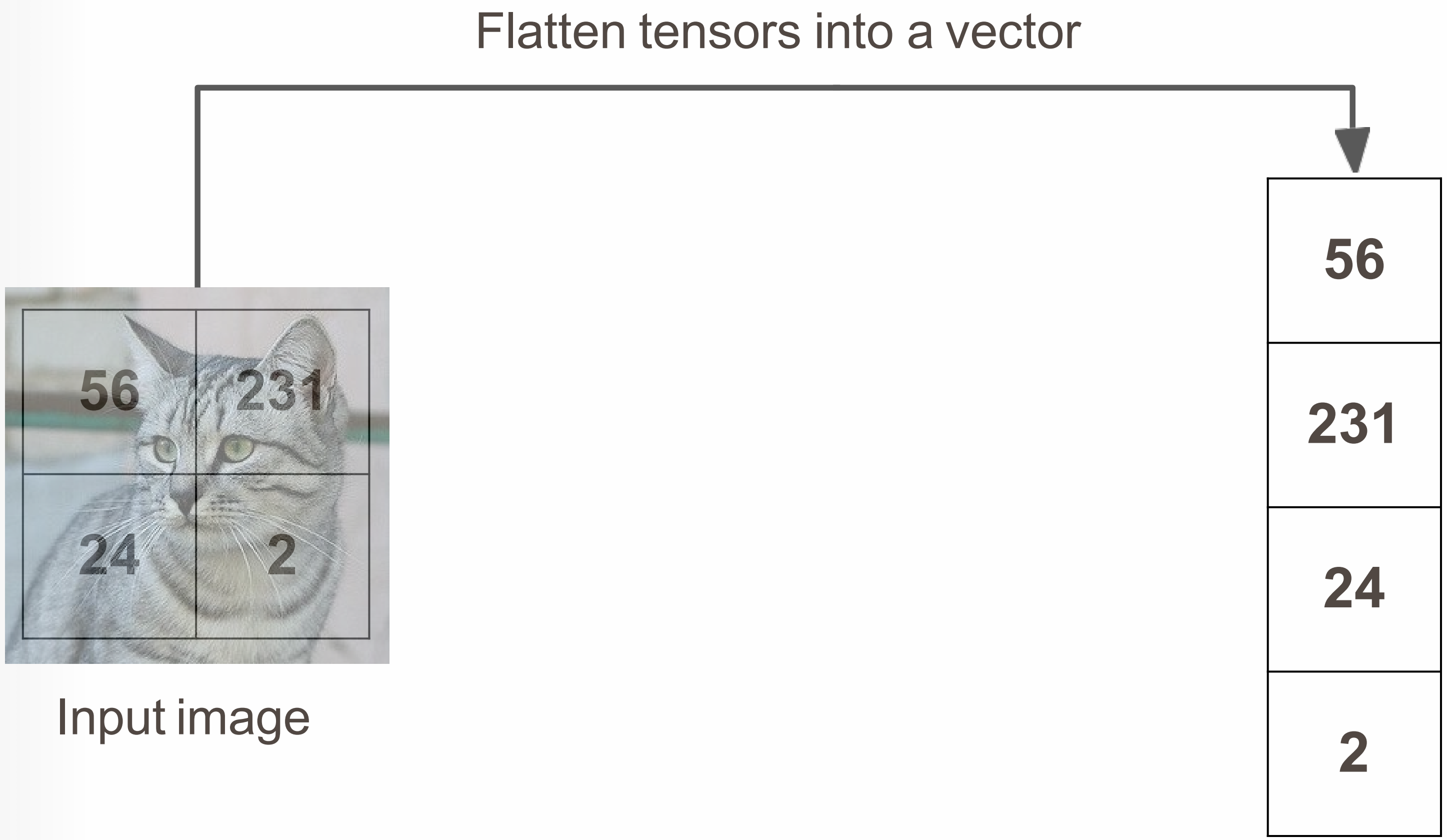
Recall CIFAR10

airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck

50,000 training images
each image is **32x32x3**

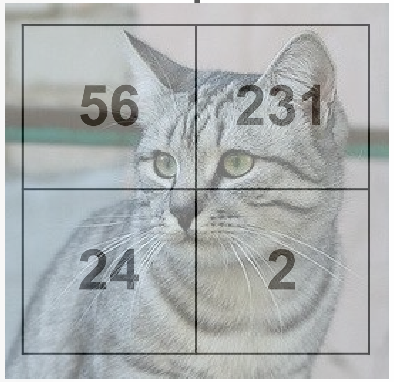
10,000 test images.

Example with an image with 4 pixels, and 3 classes (cat/dog/ship)



Input image

Example with an image with 4 pixels, and 3 classes (cat/dog/ship)

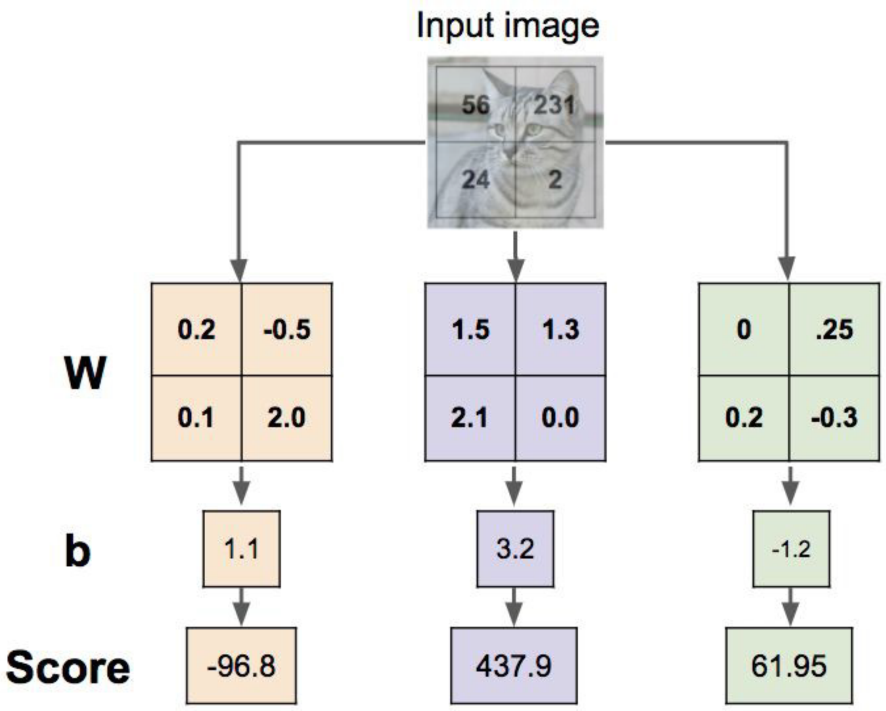


Input image

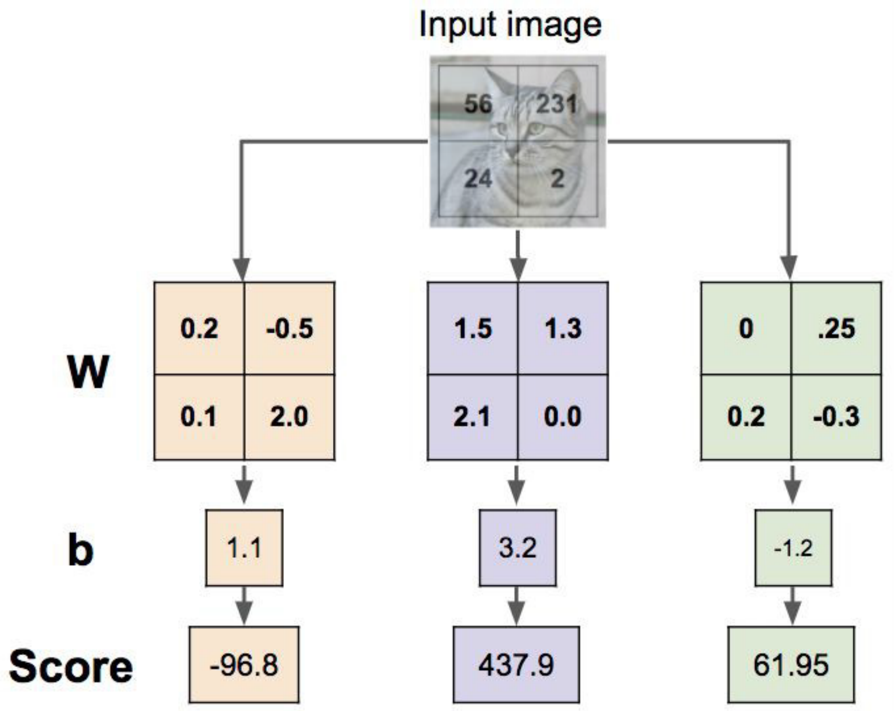
Flatten tensors into a vector

$$\underbrace{\begin{bmatrix}0.2&-0.5&0.1&2.0\\1.5&1.3&2.1&0.0\\0&0.25&0.2&-0.3\end{bmatrix}}_{W}\begin{bmatrix}56\\231\\24\\2\end{bmatrix}+\underbrace{\begin{bmatrix}1.1\\3.2\\-1.2\end{bmatrix}}_{b}=\begin{bmatrix}-96.8\\437.9\\61.95\end{bmatrix}\begin{array}{l}\text{Cat score}\\\text{Dog score}\\\text{Ship score}\end{array}$$

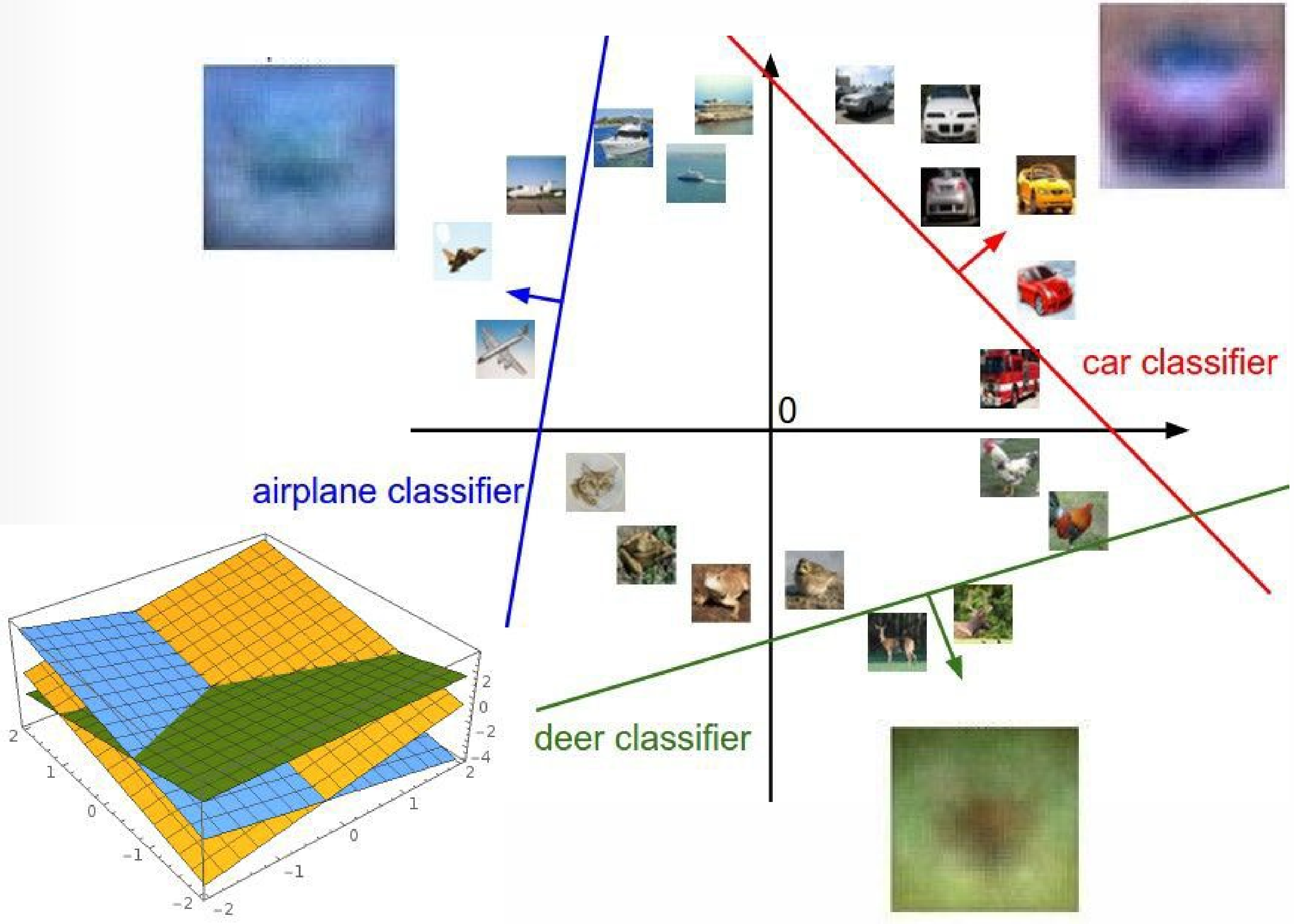
Interpreting a Linear Classifier



Interpreting a Linear Classifier: Visual Viewpoint



Interpreting a Linear Classifier: Geometric Viewpoint



$$f(x,W) = Wx + b$$

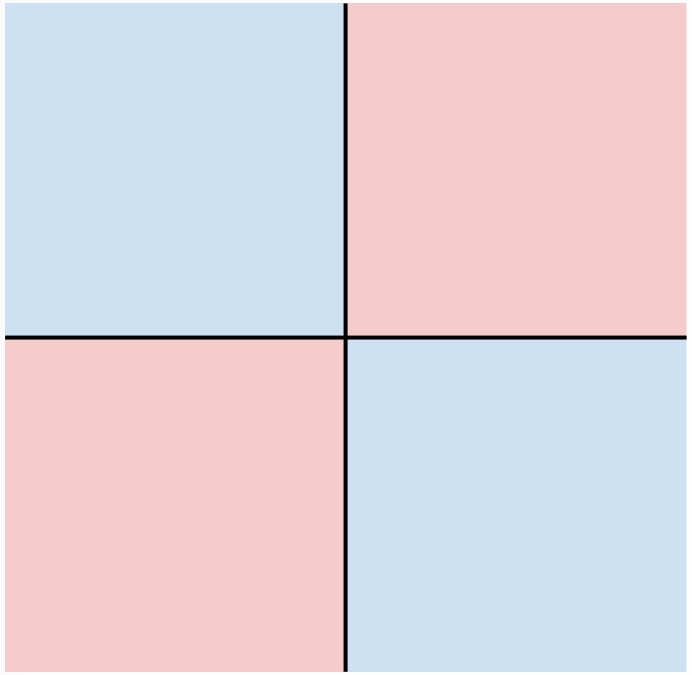
Array of **32x32x3** numbers
(3072 numbers total)

Plot created using Wolfram Cloud

Hard cases for a linear classifier

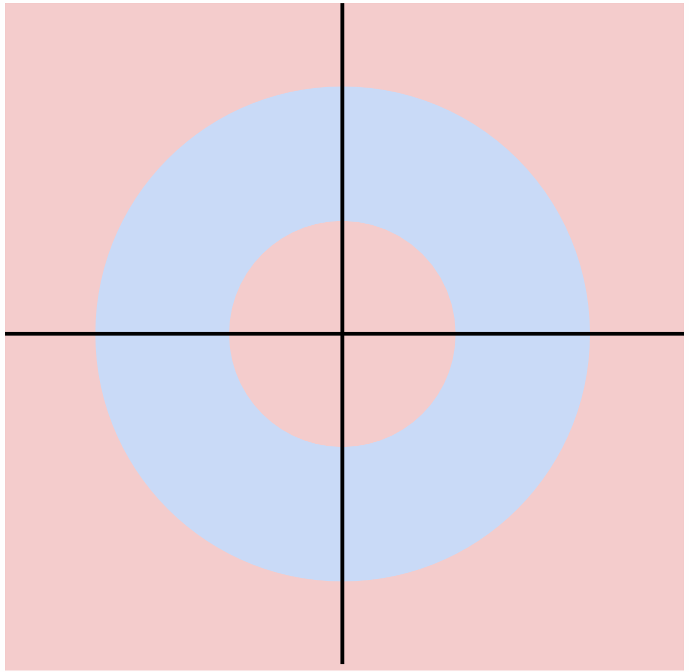
Class 1:
First and third quadrants

Class 2:
Second and fourth quadrants



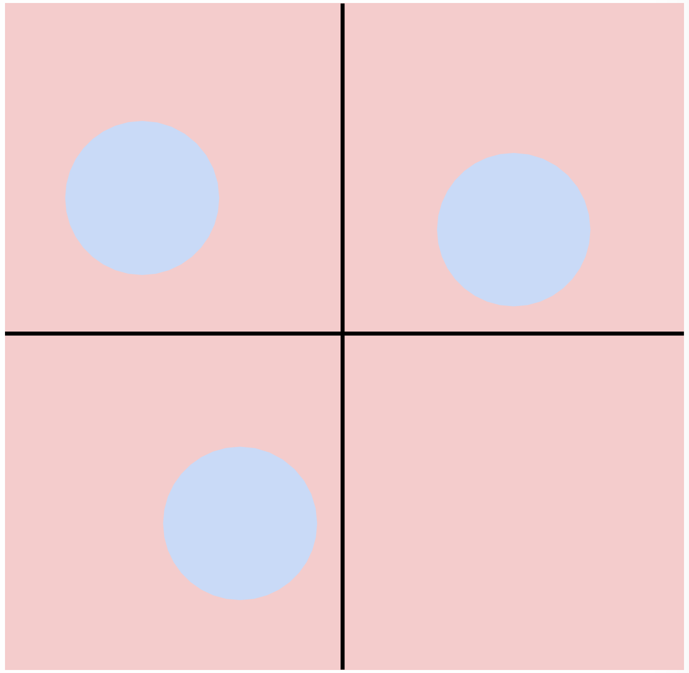
Class 1:
1 <= L2 norm <= 2

Class 2:
Everything else



Class 1:
Three modes

Class 2:
Everything else



Linear Classifier – Choose a good W



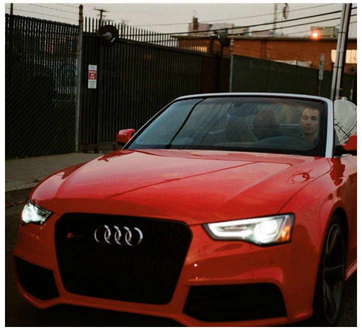
airplane	-3.45	-0.51	3.42
automobile	-8.87	**6.04**	4.64
bird	0.09	5.31	2.65
cat	**2.9**	-4.22	5.1
deer	4.48	-4.19	2.64
dog	8.02	3.58	5.55
frog	3.78	4.49	**-4.34**
horse	1.06	-4.37	-1.5
ship	-0.36	-2.09	-4.79
truck	-0.72	-2.93	6.14

TODO:

1. Define a **loss function** that quantifies our unhappiness with the scores across the training data.

2.Come up with a way of efficiently finding the parameters that minimize the loss function. **(optimization)**

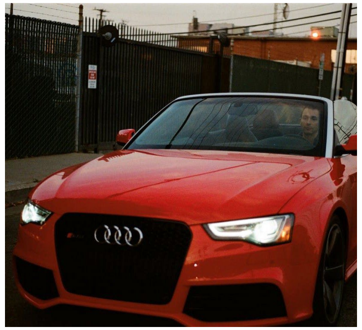
Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**

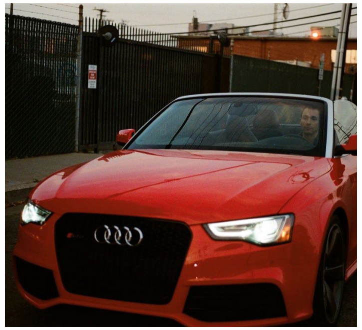
Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:

A **loss function** tells how good our current classifier is


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:

cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**

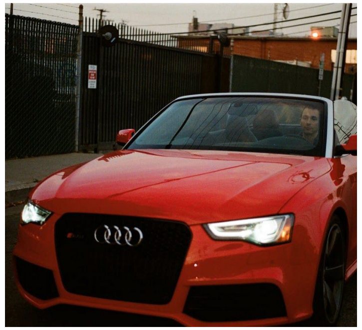
A **loss function** tells how good our current classifier is

Given a dataset of examples

$$\{(x_i, y_i)\}_{i=1}^{N}$$

Where x_i is image and y_i is (integer) label

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**

A **loss function** tells how good our current classifier is

Given a dataset of examples

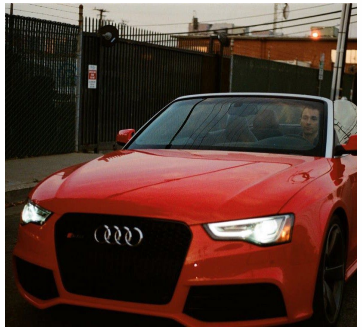
$$\{(x_i, y_i)\}_{i=1}^{N}$$

Where x_i is image and y_i is (integer) label

Loss over the dataset is a average of loss over examples:

$$L = \frac{1}{N} \sum_i L_i(f(x_i, W), y_i)$$

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**

Multiclass SVM loss:

Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

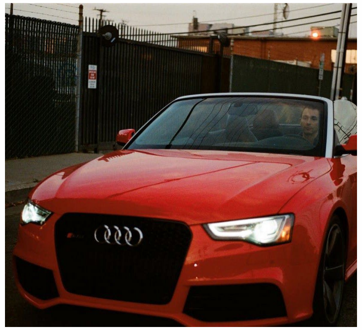
and using the shorthand for the scores vector: $s = f(x_i, W)$

the SVM loss has the form:

$$L_i = \sum_{j \neq y_i} \begin{cases} 0 & \text{if } s_{y_i} \geq s_j + 1 \\ s_j - s_{y_i} + 1 & \text{otherwise} \end{cases}$$

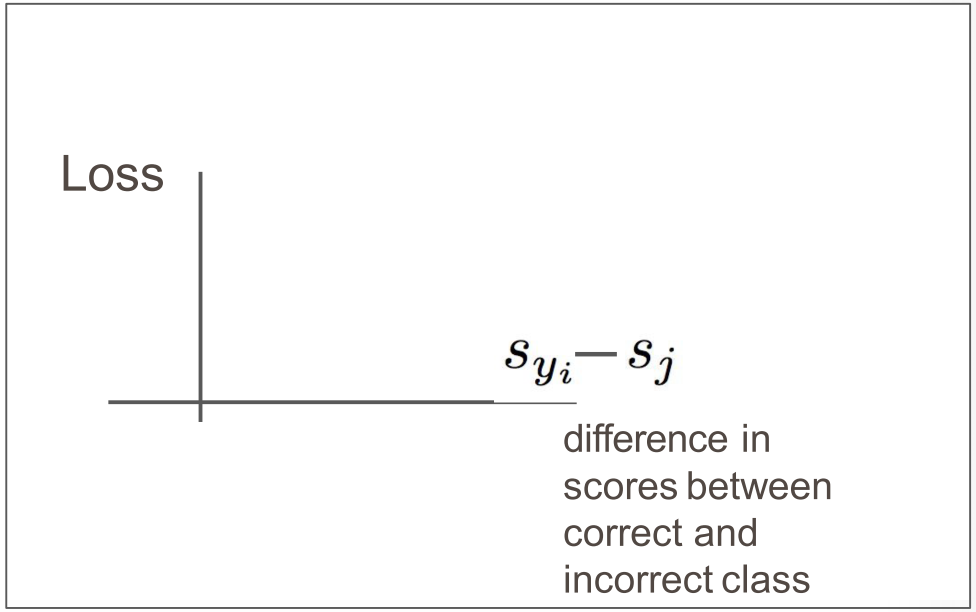
$$= \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x,W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**

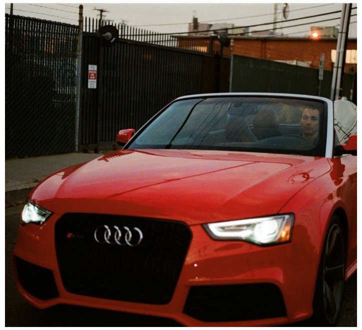
Interpreting Multiclass SVM loss:



$$L_i = \sum_{j \neq y_i} \begin{cases} 0 & \text{if } s_{y_i} \geq s_j + 1 \\ s_j - s_{y_i} + 1 & \text{otherwise} \end{cases}$$

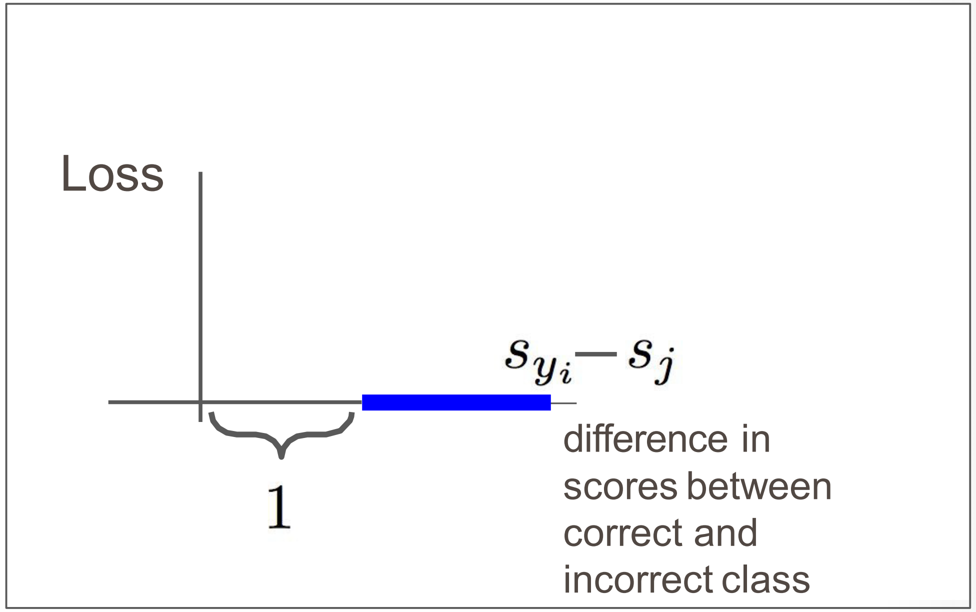
$$= \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**

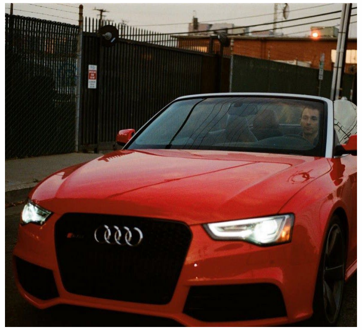
Interpreting Multiclass SVM loss:



$$L_i = \sum_{j \neq y_i} \begin{cases} 0 & \text{if } s_{y_i} \geq s_j + 1 \\ s_j - s_{y_i} + 1 & \text{otherwise} \end{cases}$$

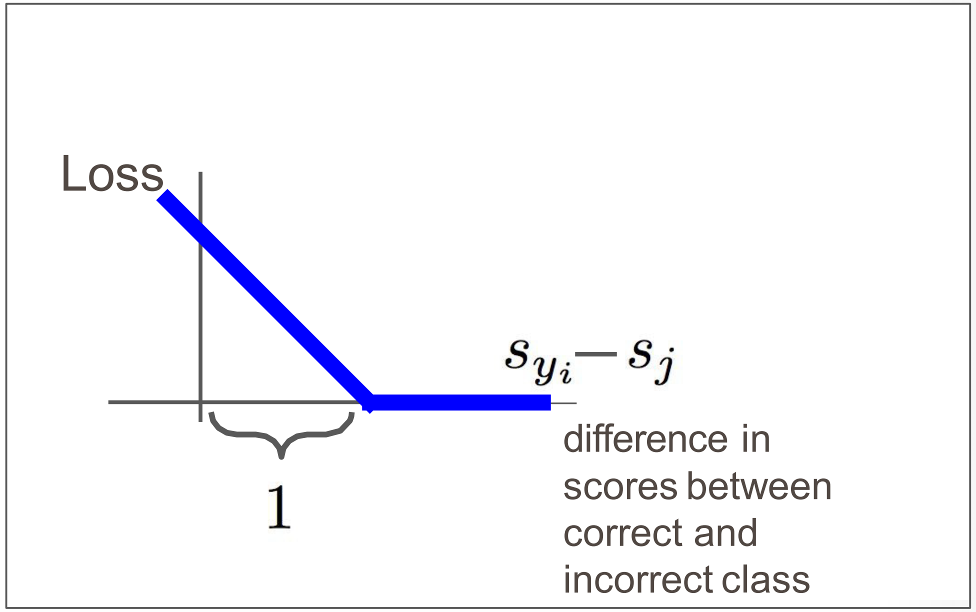
$$= \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**

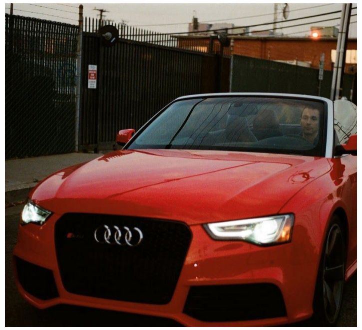
Interpreting Multiclass SVM loss:



$$L_i = \sum_{j \neq y_i} \begin{cases} 0 & \text{if } s_{y_i} \geq s_j + 1 \\ s_j - s_{y_i} + 1 & \text{otherwise} \end{cases}$$

$$= \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**

Multiclass SVM loss:

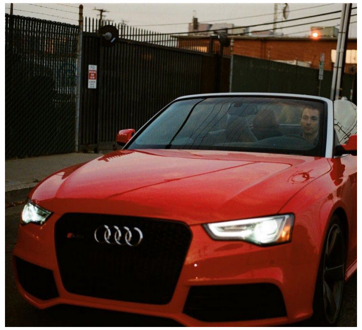
Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

and using the shorthand for the scores vector: $s = f(x_i, W)$

the SVM loss has the form:

$$\boxed{L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)}$$

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x,W) = Wx$ are:

cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**
Losses:	2.9		

Multiclass SVM loss:

Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

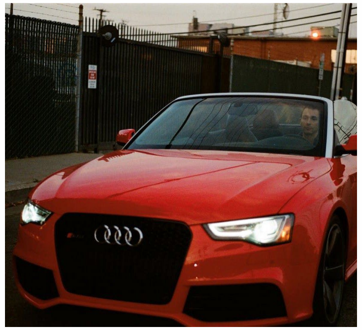
and using the shorthand for the scores vector: $s = f(x_i, W)$

the SVM loss has the form:

$$\boxed{L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)}$$

$= \max(0, 5.1 - 3.2 + 1)$
$\quad + \max(0, -1.7 - 3.2 + 1)$
$= \max(0, 2.9) + \max(0, -3.9)$
$= 2.9 + 0$
$= 2.9$

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**
Losses:	2.9	0	

Multiclass SVM loss:

Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

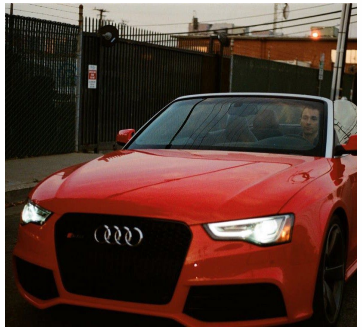
and using the shorthand for the scores vector: $s = f(x_i, W)$

the SVM loss has the form:

$$\boxed{L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)}$$

= max(0, 1.3 - 4.9 + 1)
+max(0, 2.0 - 4.9 + 1)
= max(0, -2.6) + max(0, -1.9)
= 0 + 0
= 0

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**
Losses:	2.9	0	12.9

Multiclass SVM loss:

Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

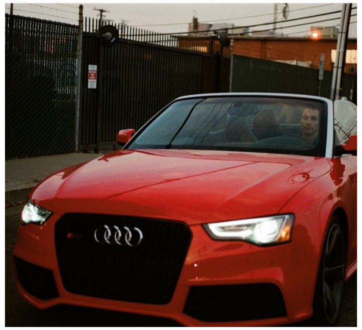
and using the shorthand for the scores vector: $s = f(x_i, W)$

the SVM loss has the form:

$$\boxed{L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)}$$

= max(0, 2.2 - (-3.1) + 1)
+max(0, 2.5 - (-3.1) + 1)
= max(0, 6.3) + max(0, 6.6)
= 6.3 + 6.6
= 12.9

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:

cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**
Losses:	2.9	0	12.9

Multiclass SVM loss:

Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

and using the shorthand for the scores vector: $s = f(x_i, W)$

the SVM loss has the form:

$$\boxed{L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)}$$

Loss over full dataset is average:

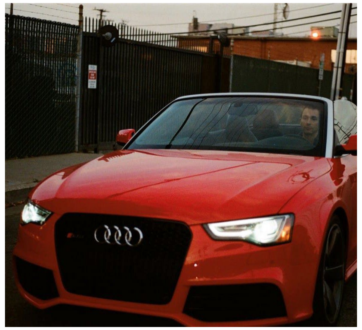
$$L = \frac{1}{N} \sum_{i=1}^{N} L_i$$

L = (2.9 + 0 + 12.9)/3
= **5.27**

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:

Multiclass SVM loss:

$$\boxed{L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)}$$



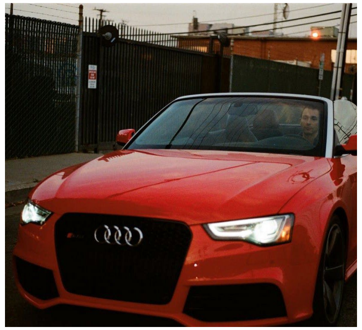
cat	1.3
car	**4.9**
frog	2.0
Losses:	0

Q1: What happens to loss if car scores decrease by 0.5 for this training example?

Q2: what is the min/max possible SVM loss L_i?

Q3: At initialization W is small so all s ≈ 0. What is the loss L_i, assuming N examples and C classes?

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**
Losses:	2.9	0	12.9

Multiclass SVM loss:

Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

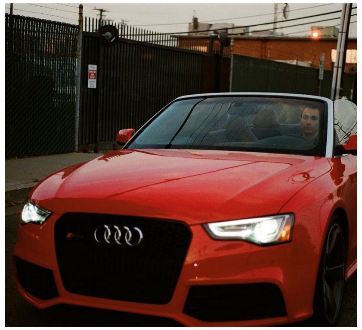
and using the shorthand for the
scores vector: $s = f(x_i, W)$

the SVM loss has the form:

$$\boxed{L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)}$$

Q4: What if the sum was over all classes? (including j = y_i)

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**
Losses:	2.9	0	12.9

Multiclass SVM loss:

Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

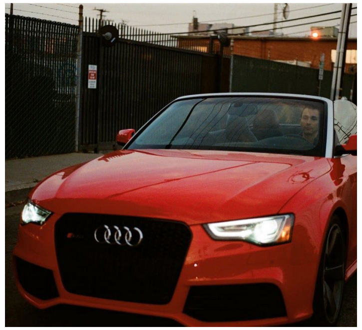
and using the shorthand for the scores vector: $s = f(x_i, W)$

the SVM loss has the form:

$$\boxed{L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)}$$

Q5: What if we used mean instead of sum?

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:

cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**
Losses:	2.9	0	12.9

Multiclass SVM loss:

Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

and using the shorthand for the scores vector: $s = f(x_i, W)$

the SVM loss has the form:

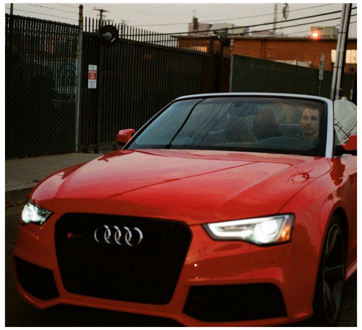
$$\boxed{L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)}$$

Q6: What if we used

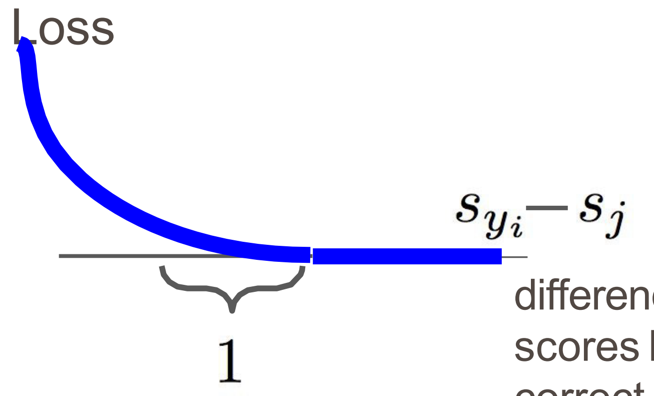
$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)^2$$

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:

Multiclass SVM loss:


cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**
Losses:	2.9	0	12.9



Q6: What if we used

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)^2$$

Multiclass SVM Loss: Example code

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

```
def L_i_vectorized(x, y, W):
  scores = W.dot(x)
  margins = np.maximum(0, scores - scores[y] + 1)
  margins[y] = 0
  loss_i = np.sum(margins)
  return loss_i
```

First calculate scores
Then calculate the margins $s_j - s_{yi} + 1$
only sum j is not y_i, so when $j = y_i$, set to zero. # sum across all j

SOFTMAX CLASSIFIER

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

cat	**3.2**
car	5.1
frog	-1.7

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$\boxed{s = f(x_i; W)}$$

$$\boxed{P(Y = k|X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}}$$

Softmax Function

cat	**3.2**
car	5.1
frog	-1.7

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$s = f(x_i; W)$$

$$P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}$$

Softmax Function

Probabilities must be >= 0

		$\xrightarrow{\text{exp}}$	
cat	**3.2**		**24.5**
car	5.1		164.0
frog	-1.7		0.18

unnormalized probabilities

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$s = f(x_i; W)$$

$$P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}$$

Softmax Function

Probabilities must be >= 0

Probabilities must sum to 1

		exp		normalize	
cat	**3.2**		**24.5**		**0.13**
car	5.1	→	164.0	→	0.87
frog	-1.7		0.18		0.00

unnormalized probabilities

probabilities

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$s = f(x_i; W)$$

$$P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}$$

Softmax Function

Probabilities must be >= 0

Probabilities must sum to 1

	Unnormalized log-probabilities / logits	$\xrightarrow{\text{exp}}$	unnormalized probabilities	$\xrightarrow{\text{normalize}}$	probabilities
cat	**3.2**		**24.5**		**0.13**
car	5.1		164.0		0.87
frog	-1.7		0.18		0.00

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$s = f(x_i; W)$$

$$P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}$$ Softmax Function

$$L_i = -\log P(Y = y_i | X = x_i)$$

Probabilities must be >= 0

Probabilities must sum to 1

	Unnormalized log-probabilities / logits		unnormalized probabilities		probabilities
cat	**3.2**	exp →	**24.5**	normalize →	**0.13**
car	5.1		164.0		0.87
frog	-1.7		0.18		0.00

→ $L_i = -\log(0.13)$
$= \mathbf{2.04}$

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$s = f(x_i; W)$$

$$P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}$$

Softmax Function

Probabilities must be >= 0

Probabilities must sum to 1

$$L_i = -\log P(Y = y_i | X = x_i)$$

	Unnormalized log-probabilities / logits	$\xrightarrow{\text{exp}}$ unnormalized probabilities	$\xrightarrow{\text{normalize}}$ probabilities
cat	**3.2**	**24.5**	**0.13**
car	5.1	164.0	0.87
frog	-1.7	0.18	0.00

$\rightarrow$ $L_i = -\log(0.13)$ $= \mathbf{2.04}$

Maximum Likelihood Estimation
Choose weights to maximize the likelihood of the observed data
(See CS 229 for details)

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$s = f(x_i; W)$$

$$P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}$$

Softmax Function

Probabilities must be >= 0

Probabilities must sum to 1

$$L_i = -\log P(Y = y_i | X = x_i)$$

	Unnormalized log-probabilities / logits	exp → unnormalized probabilities	normalize → probabilities	compare	Correct probs
cat	**3.2**	**24.5**	**0.13**		**1.00**
car	5.1	164.0	0.87		0.00
frog	-1.7	0.18	0.00		0.00

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$s = f(x_i; W)$$

$$P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}$$

Softmax Function

Probabilities must be >= 0

Probabilities must sum to 1

$$L_i = -\log P(Y = y_i | X = x_i)$$

	Unnormalized log-probabilities / logits	$\xrightarrow{\text{exp}}$	unnormalized probabilities	$\xrightarrow{\text{normalize}}$	probabilities		Correct probs
cat	**3.2**		**24.5**		**0.13**	→ compare ←	**1.00**
car	5.1		164.0		0.87		0.00
frog	-1.7		0.18		0.00		0.00

Kullback–Leibl r divergence

$$D_{KL}(P \| Q) = \sum_y P(y) \log \frac{P(y)}{Q(y)}$$

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$s = f(x_i; W)$$

$$P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}$$

Softmax Function

Probabilities must be >= 0

Probabilities must sum to 1

$$L_i = -\log P(Y = y_i | X = x_i)$$

	Unnormalized log-probabilities / logits		unnormalized probabilities		probabilities		Correct probs
cat	**3.2**	exp →	**24.5**	normalize →	**0.13**	→ compare ←	**1.00**
car	5.1		164.0		0.87		0.00
frog	-1.7		0.18		0.00		0.00

Cross Entropy

$$H(P, Q) = H(p) + D_{KL}(P \| Q)$$

Softmax Classifier (Multinomial Logistic Regression)

Want to interpret raw classifier scores as **probabilities**

$$\boxed{s = f(x_i; W)}$$

$$\boxed{P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}}$$

Softmax Function

Maximize probability of correct class

$$L_i = -\log P(Y = y_i | X = x_i)$$

Putting it all together:

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

cat	**3.2**
car	5.1
frog	-1.7

Softmax Classifier (Multinomial Logistic Regression)

cat	**3.2**
car	5.1
frog	-1.7

Want to interpret raw classifier scores as **probabilities**

$$\boxed{s = f(x_i; W)}$$

$$\boxed{P(Y = k|X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}}$$

Softmax Function

Maximize probability of correct class

$$L_i = -\log P(Y = y_i|X = x_i)$$

Putting it all together:

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

Q1: What is the min/max possible softmax loss L_i?

Q2: At initialization all s_j will be approximately equal; what is the softmax loss L_i, assuming C classes?

Softmax Classifier (Multinomial Logistic Regression)

cat 3.2

car 5.1

frog -1.7

Want to interpret raw classifier scores as **probabilities**

$$s = f(x_i; W)$$

$$P(Y = k | X = x_i) = \frac{e^{s_k}}{\sum_j e^{s_j}}$$

Softmax Function

Maximize probability of correct class

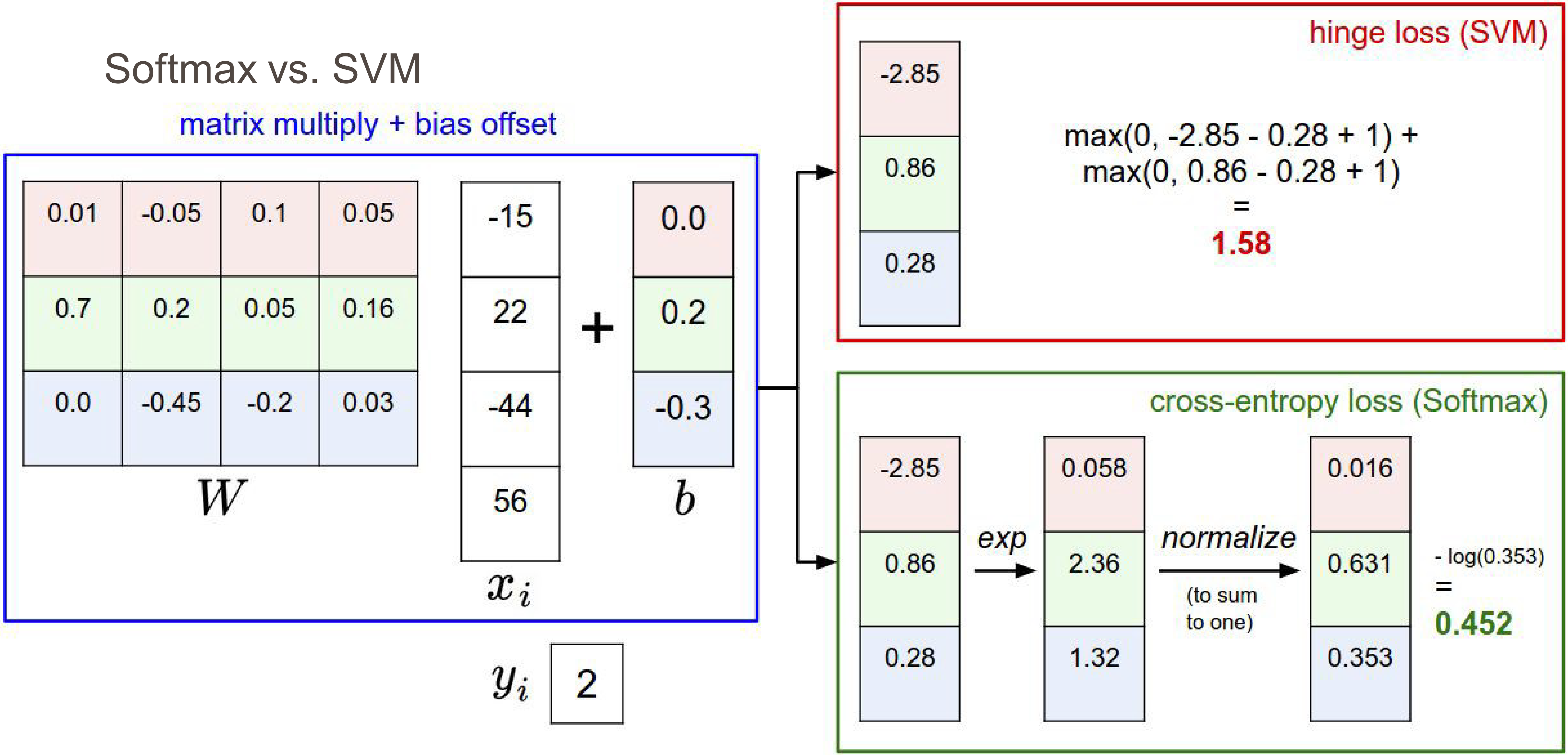
$$L_i = -\log P(Y = y_i | X = x_i)$$

Putting it all together:

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

Q2: At initialization all s will be approximately equal; what is the loss?
A: -log(1/C) = log(C),
If C = 10, then L_i = log(10) ≈ 2.3

Softmax vs. SVM



Softmax vs. SVM

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

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

Softmax vs. SVM

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

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

assume scores:

[10, -2, 3]

[10, 9, 9]

[10, -100, -100]

and $\boxed{y_i = 0}$

Q: What is the **softmax loss** and the **SVM** loss?

Softmax vs. SVM

$$L_i = -\log\left(\frac{e^{s_{y_i}}}{\sum_j e^{s_j}}\right) \qquad L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

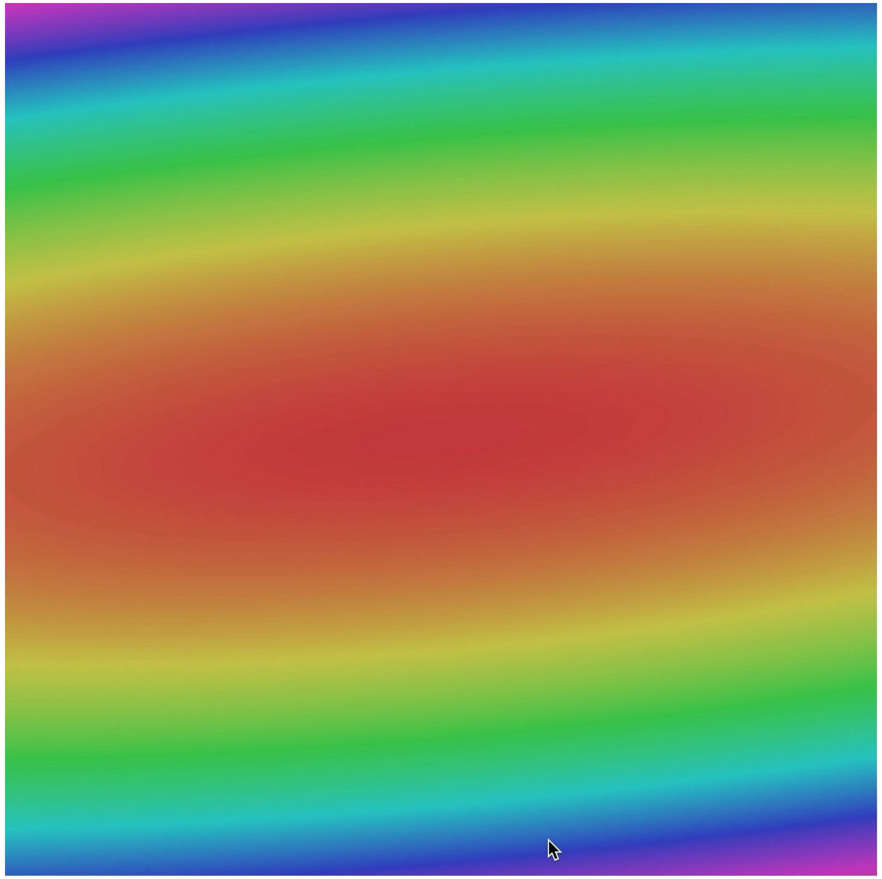
assume scores:
[20, -2, 3]
[20, 9, 9]
[20, -100, -100]
and $\boxed{y_i = 0}$

Q: What is the **softmax loss** and the **SVM** loss **if I double the correct class score from 10 -> 20**?

Coming up:

- Regularization
- Optimization

$$f(x,W) = Wx + b$$

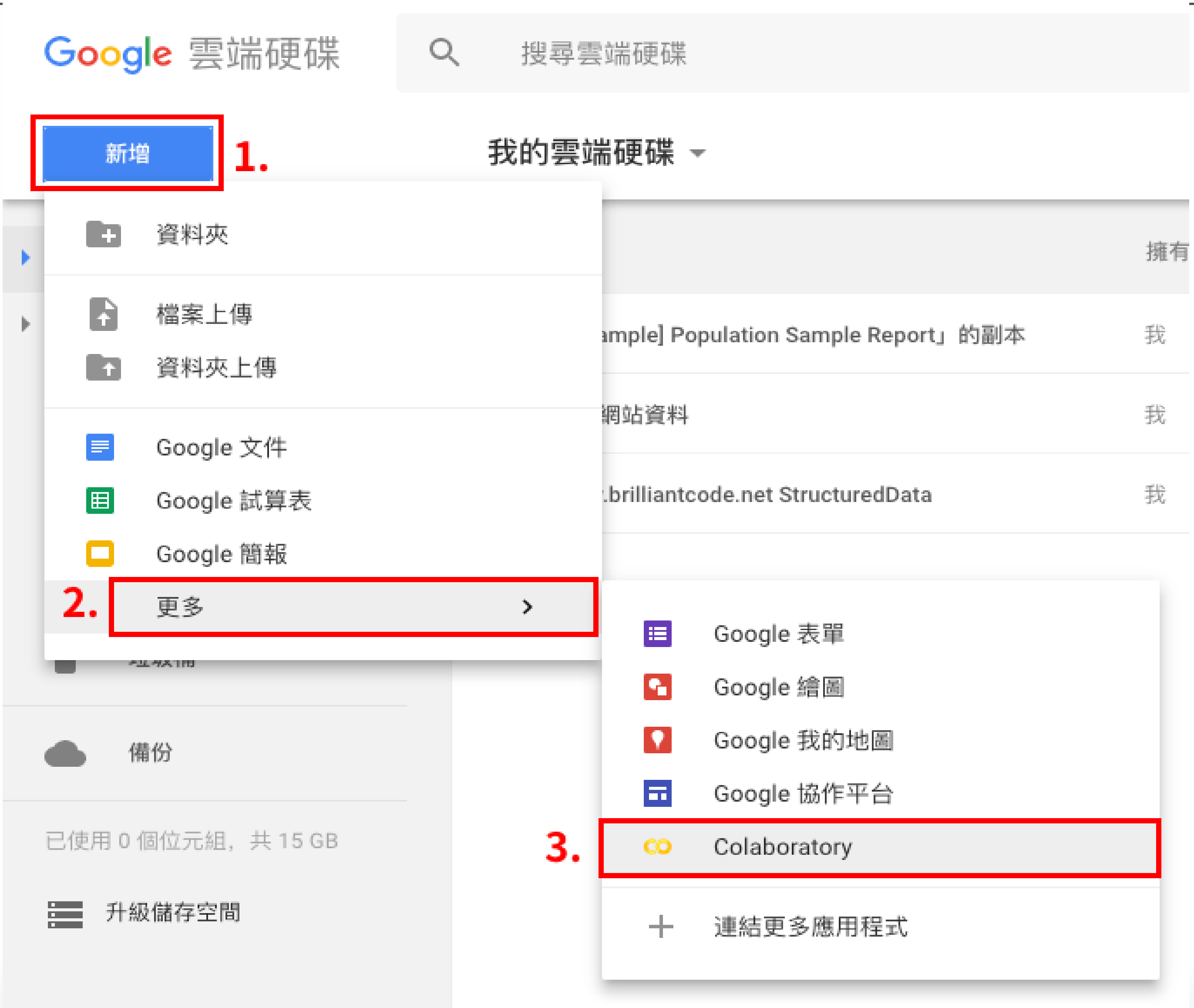


GOOGLE CLOUD RESOURCE: COLAB

Run TF without GPU?

- Google COLAB (Colaboratory)
  - Only three steps
    - Having free GPU/TPU resource if you have a google account

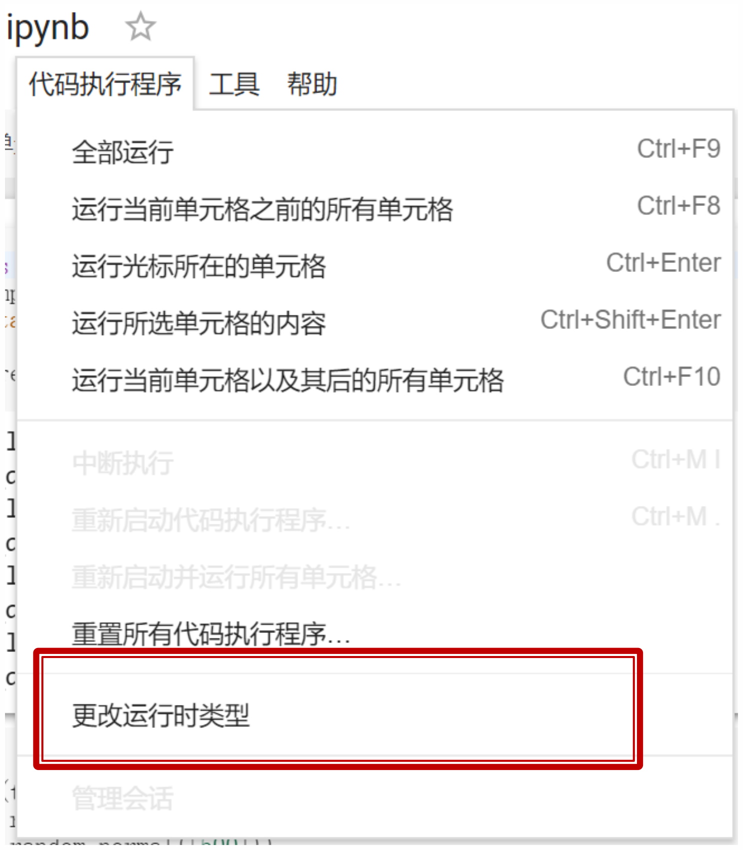
Click right key on ipynb and find colab



COLAB

- Free GPU (K80/V100) and TPU resource
  - 24G RAM (very large, compared to standard GPU)
    - RTX 2080 Ti
      - NTD 35,000
    - K80
      - NTD 180, 000
- You only can run a program on the free GPU 12 hours per day
  - Wait another 12 hours to access the free resource
- Runtime environment
  - Can be CPU/GPU/TPU

GPU Resource Usage



GPU vs TPU? TPU 為 Google 自家提出的加速硬體，照道理應該會比較快! 大家亦可試試看。

Check Whether The GPU Resource is Enabled

```
import tensorflow as tf
device_name = tf.test.gpu_device_name()
if device_name != '/device:GPU:0':
    raise SystemError('GPU device not found')
print('Found GPU at: {}'.format(device_name))
```

- 若有使用到GPU，則會顯示

```
import tensorflow as tf
device_name = tf.test.gpu_device_name()
if device_name != '/device:GPU:0':
  raise SystemError('GPU device not found')
print('Found GPU at: {}'.format(device_name))
```

```
Found GPU at: /device:GPU:0
```

COLAB with Your Own Files

```
from google.colab import drive
import os

drive.mount('/content/gdrive'  )
os.chdir("/content/gdrive/My Drive") #更改路徑
os.getcwd() #查看當前路徑

Use ! To run bash shell in colab
!ls
## show your files
```

- 將上述程式碼貼到 COLAB 中，並根據【畫面指示】操作，就可以有與 GDRIVE連線的權限

Connect Google Drive with COLAB

```
!mkdir -p Drive
!google-drive-ocamlfuse Drive
```

- 上述指令在於
  - 在 COLAB 中建立一個資料夾 Drive (遠端中)
  - 利用指令!google-drive-ocamlfuse，把你連結到的雲端硬碟，連結到 Drive 資料夾中
  - 換句話說，你存取 Drive 資料夾，等於存取你的雲端硬碟

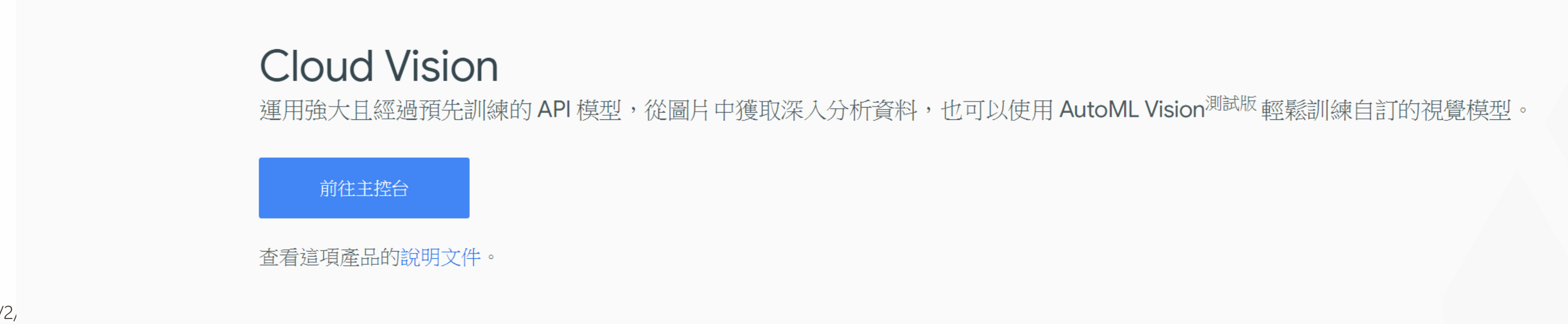
```
!ls Drive
```

```
 00125870.pdf
 00125870.pdf.odt
'00125870.pdf 的翻譯版本.odt'
'1030909-清大(電機工程學系 ) SERVER.xlsx'
'1030909-清大(電機工程學系 ) SERVER.xlsx.ods'
'1030910-清大(電機工程學系 ) SERVER.pdf'
```

GOOGLE CLOUD SERVICE: VISION (AUTOML)

線上免費模型 (12個月免費)

- https://cloud.google.com/vision



AutoML 授權申請 (尚未有付費帳號)

免費試用 **Google Cloud Platform**

步驟 2 之 1

國家/地區

台灣

服務條款

☐ 我同意《Google Cloud Platform 服務條款》和所有適用服務和 API 的服務條款。我也已詳閱並同意遵守《Google Cloud Platform 免費試用版的服務條款》。

必須勾選這個核取方塊才能提交表單

電子郵件最新消息

☐ 我想要定期收到 Google Cloud 和 Google Cloud Partner 傳送的電子郵件，隨時掌握相關新聞、產品動態和特價優惠。

同意並繼續

使用所有 **Cloud Platform** 產品

取得建構及執行應用程式、網站和服務所需的一切資源，包括 Firebase 和 Google Maps API。

免費獲得 **$300** 美元的試用額度

申請試用即可獲得 $300 美元的試用額度，可於未來 12 個月內在 Google Cloud Platform 上使用。

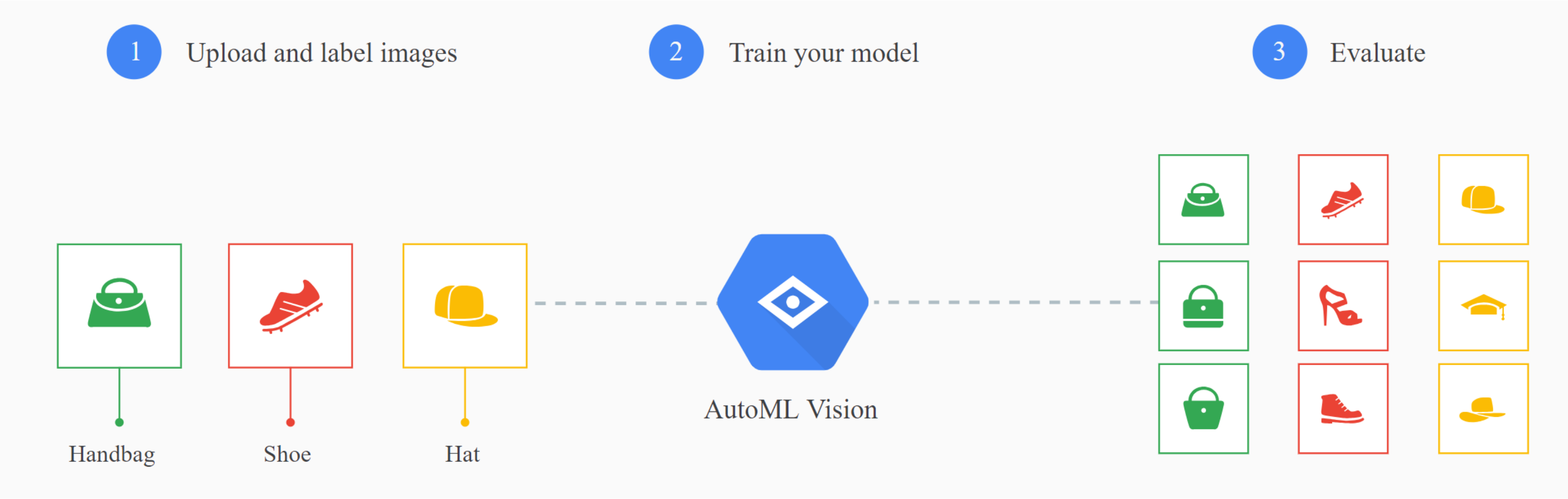
免費試用期結束後不會自動向您收費

我們之所以要求您提供信用卡資訊，是為了確保您不是自動化程式。除非您手動升級至付費帳戶，否則我們不會向您收費。

AutoML 授權申請 (尚未有付費帳號)

- 必須有信用卡
  - 不會偷偷被扣款
- 一般銀行的卡?
  - 台灣尚不支援
- 其他方法?
  - 請自行Google

Cloud Vision (AutoML) 使用方法



- 首先建立資料庫 (資料夾: 類別名稱)

名稱	修改日期	類型
5498592	2019/3/15 上午 0...	檔案資料夾
5513628	2019/3/15 上午 0...	檔案資料夾
5540978	2019/3/15 上午 0...	檔案資料夾

AutoML訓練自己資料庫方法

- 建立 Cloud Storage Bucket (CSB)
- 建立 CSB 的權限
- 上傳資料到自己的CSB
  - Including image files and their labels (.csv)
- 建立自己 CSB 的 Training dataset
- 訓練模型 & Testing

AutoML: 建立 CSB

- 首先開啟 Cloud AutoML 以及 Storage 的 API

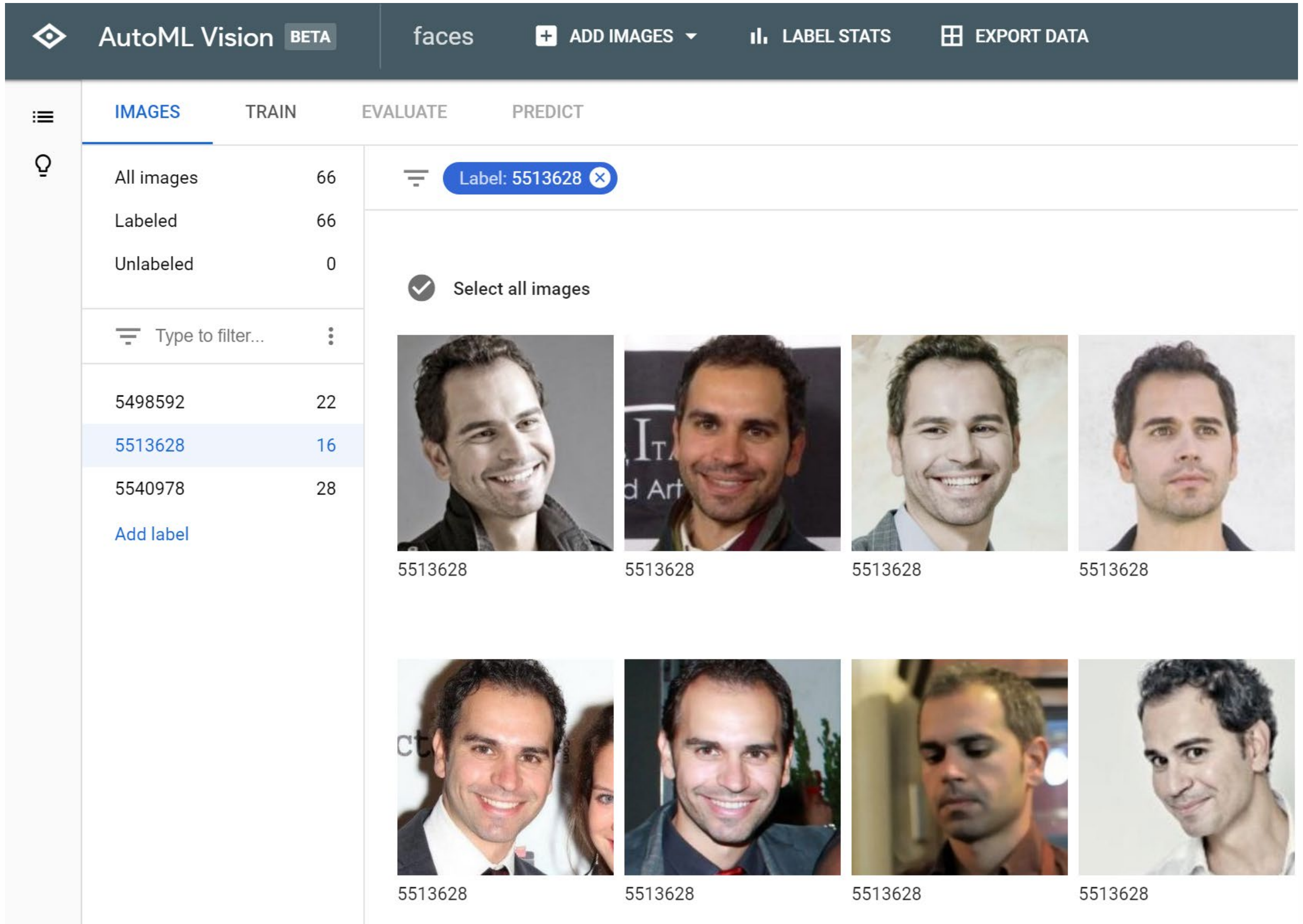
http://console.cloud.google.com/?cloudshell=true

- 首先連結到 Google cloud 主控台 Console
  - PROJECT=$(gcloud config get-value project) && BUCKET="${PROJECT}-vcm"
  - gsutil mb -p ${PROJECT} -c regional -l us-central1 gs://${BUCKET}
- 建立 Service

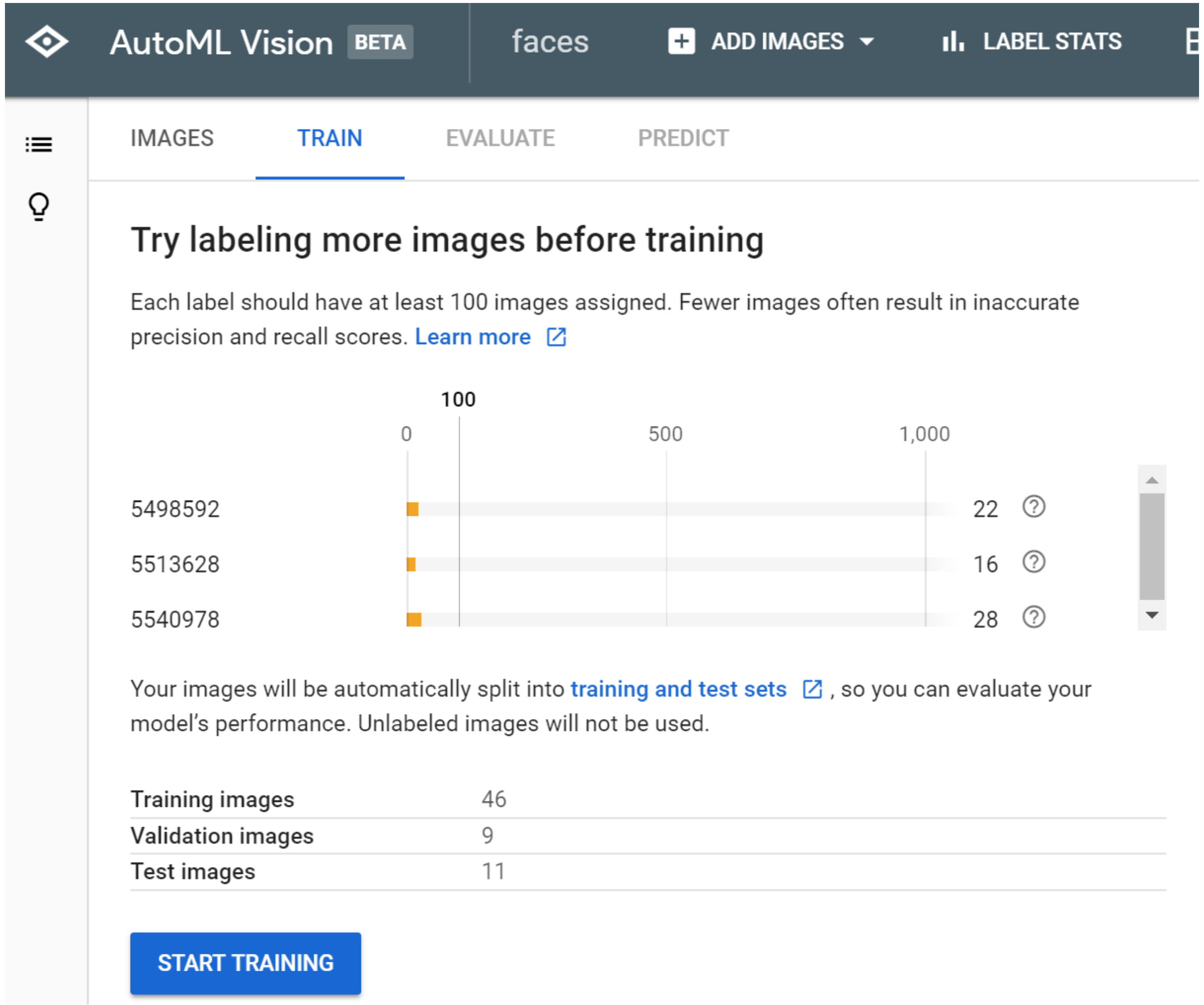
```
PROJECT=$(gcloud config get-value project)
gcloud projects add-iam-policy-binding $PROJECT \ --member="serviceAccount:custom-vision@appspot.gserviceaccount.com" \ --role="roles/ml.admin"
gcloud projects add-iam-policy-binding $PROJECT \ --member="serviceAccount:custom-vision@appspot.gserviceaccount.com" \ --role="roles/storage.admin"
```

建立自己的資料庫

- 進入到 AutoML Vision 的頁面

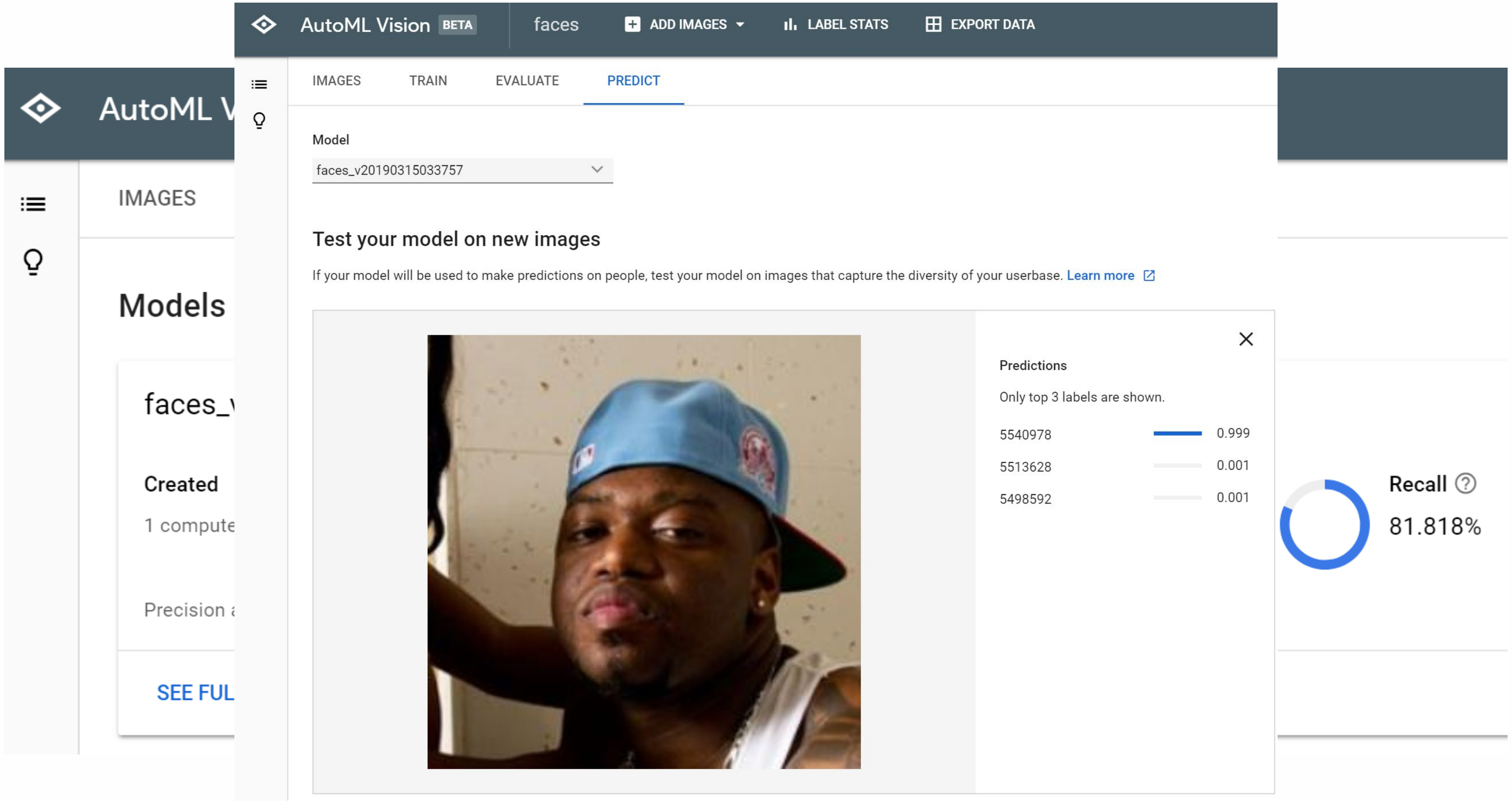


訓練MODEL


AutoML Vision BETA
faces
ADD IMAGES
LABEL STATS
IMAGES
TRAIN
EVALUATE
PREDICT
Try labeling more images before training
Each label should have at least 100 images assigned. Fewer images often result in inaccurate precision and recall scores. Learn more
100
0
500
1,000
5498592
22
5513628
16
5540978
28
Your images will be automatically split into training and test sets, so you can evaluate your model's performance. Unlabeled images will not be used.
Training images
46
Validation images
9
Test images
11
START TRAINING

Learned Model

- 可以馬上看Training 結果，Predict 可以上傳照片並辨識資料



Comparison

	Tensorflow (Own PC)	COLAB (Tensorflow)	AutoML
執行速度	快	中	中（要網路）
訓練速度	快	中	快（已上傳的話）
建立方法速度	慢	慢	超快
方便性	一般（要寫程式）	優（有網路就可）	優（要註冊，未來要錢）
客製化彈性	優（自己建立模型）	優	無
辨識效能	高	高	高一點
任務置換	容易	容易	麻煩（要換資料）
I/O效能	高（SSD版）	低（有限制）	很高
成本	一般（GPU的錢）	低（只要網路）	中（算訓練時間收費）