BASES DE DATOS EN IMAGENES

AUTOR: JIMENA OLVERES MONTIEL

1

2

3

IMAGENET Large Scale Visual Recognition Challenge

Competencia de Reconocimiento Visual

- Desde el 2010, ImageNet Large Scale Visual Recognition Challenge (ILSVRC)
- Competencia donde los equipos de investigación evalúan sus algoritmos sobre una base de datos de imágenes conocida.
- Compiten para obtener las mejores calificaciones en diversas tareas de reconocimiento visual
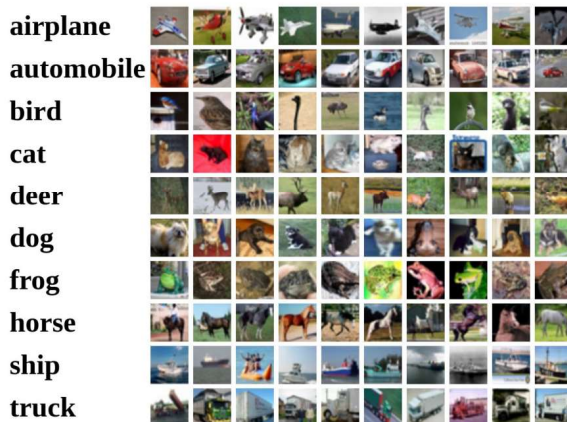
4

ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

- Tarea1: Clasificación
- Tarea 2: Clasificación con localización
- Tarea 3: Clasificacion Fina (Nueva)
  - En mas de 100+ categoría de perros.
  - Para cada categoría de perros predecir si el perro especificado dentro de una caja (bounding box) en la imagen de prueba es de una categoría particular

5

Recall CIFAR10

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

10,000 test images.

The CIFAR-10 dataset consists of 60,000 32x32 colour images in 10 classes, with 60,000 images per class. There are 50,000 training images and 10,000 test images.

The dataset is divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly-selected images from each class. The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another. Between them, the training batches contain exactly 5000 images from each class.
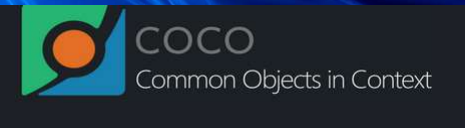
6

The CIFAR-100 dataset
This dataset is just like the CIFAR-10, except it has 100 classes containing 600 images each. There are 500 training images and 100 testing images per class. The 100 classes in the CIFAR-100 are grouped into 20 superclasses. Each image comes with a "fine" label (the class to which it belongs) and a "coarse" label (the superclass to which it belongs).

Superclass	Classes
aquatic mammals	beaver, dolphin, otter, seal, whale
fish aquarium	fish, flatfish, ray, shark, trout
flowers	orchids, poppies, roses, sunflowers, tulips
food containers	bottles, bowls, cans, cups, plates
fruit and vegetables	apples, mushrooms, oranges, pears, sweet peppers
household electrical devices	clock, computer keyboard, lamp, telephone, television
household furniture	bed, chair, couch, table, wardrobe
insects	bee, beetle, butterfly, caterpillar, cockroach
large carnivores	bear, leopard, lion, tiger, wolf

7

COCO
Common Objects in Context

- COCO is a large-scale object detection, segmentation, and captioning dataset. COCO has several features:
- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 330K images (>200K labeled)
- 1.5 million object instances
- 80 object categories
- 91 stuff categories
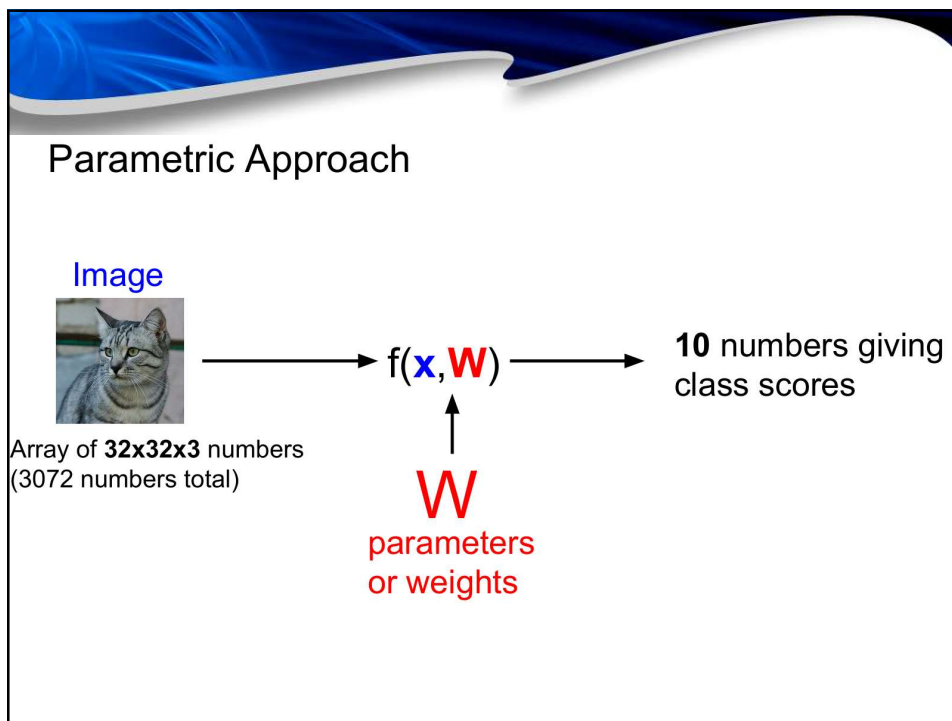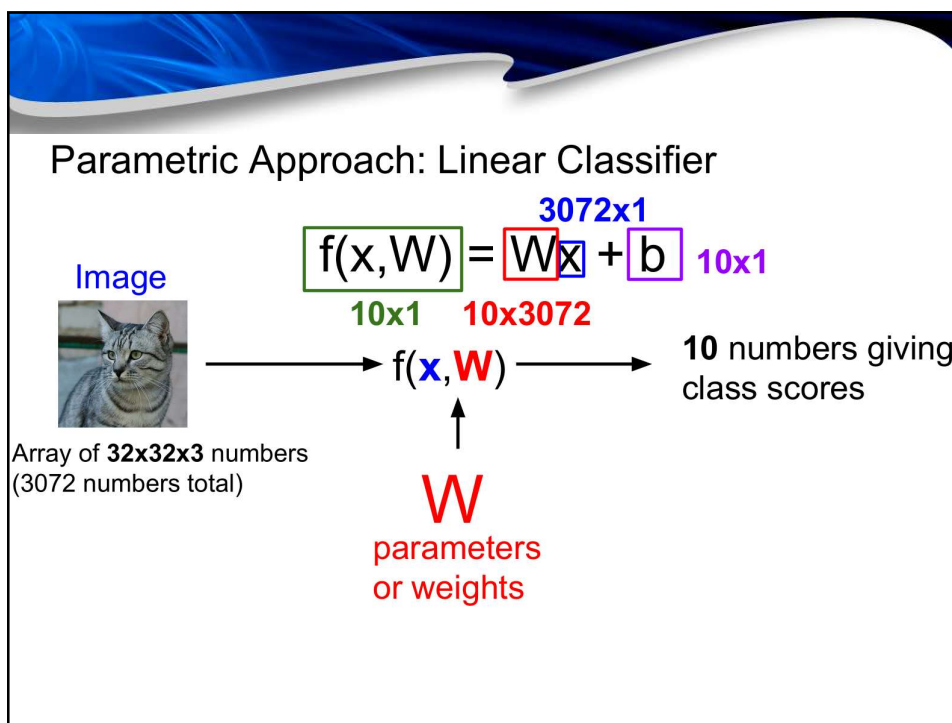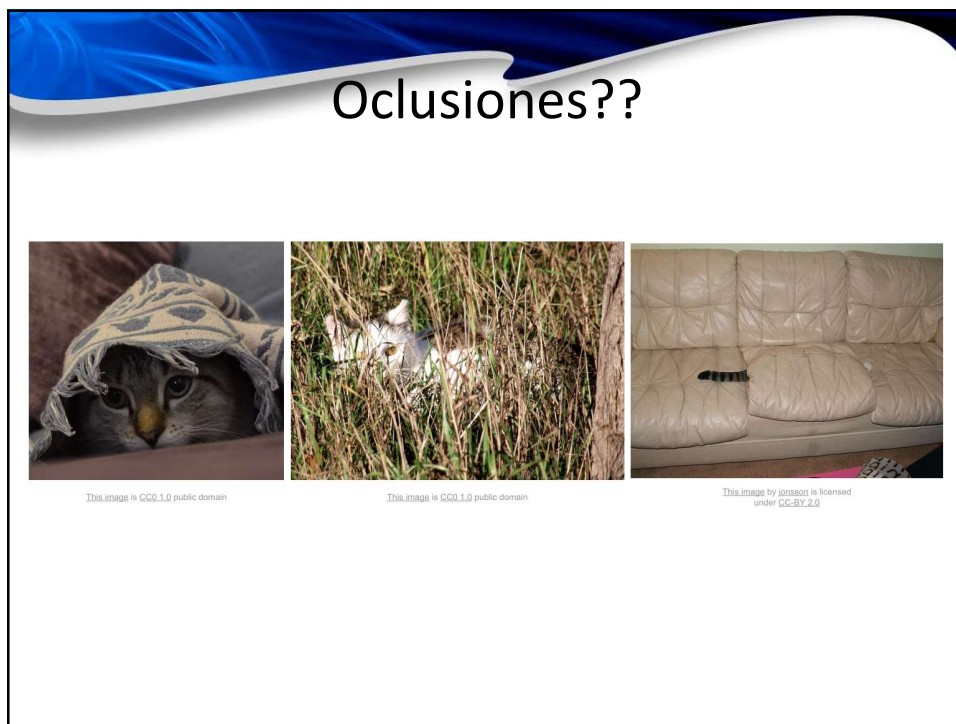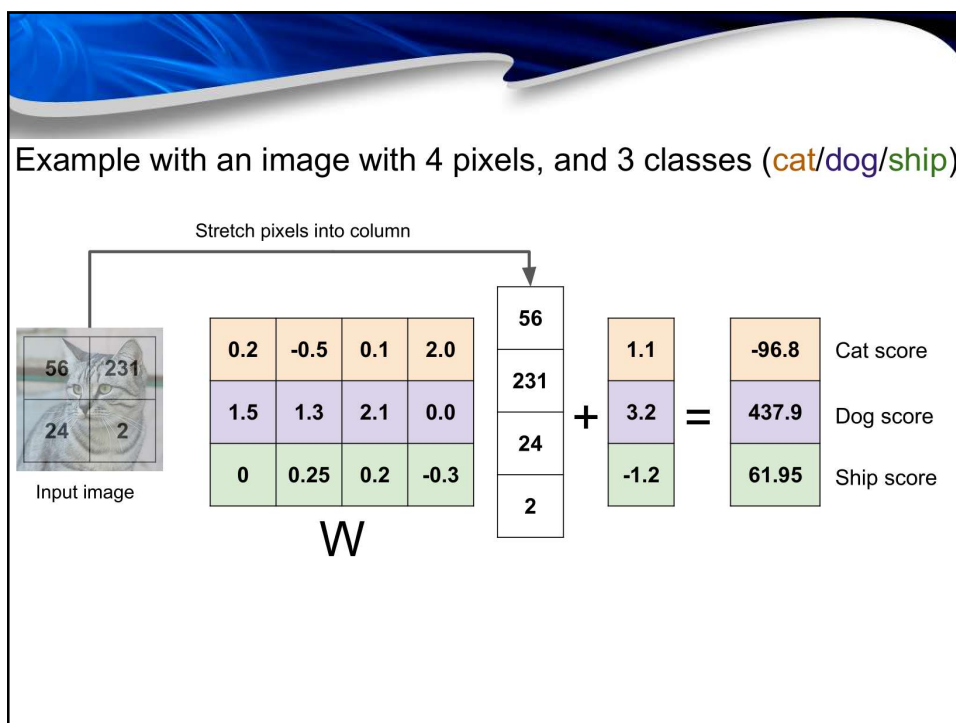- 5 captions per image

Dataset examples

8

Iris Flower
Pétalo y Sépalo

9

Aplicaciones

10

Parametric Approach

Image

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

$f(\mathbf{x},\mathbf{W})$

10 numbers giving
class scores

W

parameters
or weights

11

Parametric Approach: Linear Classifier

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

3072x1

10x1   **10x3072**   **10x1**

Image

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

$f(\mathbf{x},\mathbf{W})$

10 numbers giving
class scores

W

parameters
or weights

12

Oclusiones??

13

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

14

CIFAR 10

Interpreting a Linear Classifier
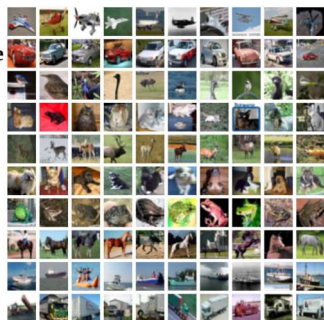
$$f(x,W) = Wx + b$$

What is this thing doing?

15

Interpreting a Linear Classifier
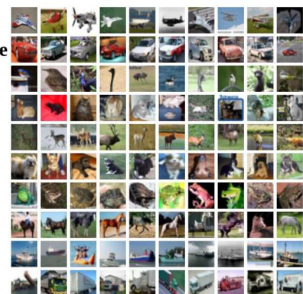
$$f(x,W) = Wx + b$$

Example trained weights of a linear classifier trained on CIFAR-10:

16

Interpreting a Linear Classifier

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

airplane classifier

car classifier

deer classifier

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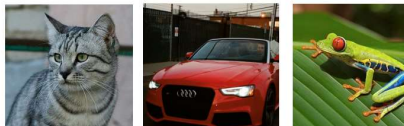(3072 numbers total)

17

Hard cases for a linear classifier

Class 1:
number of pixels > 0 odd

Class 2:
number of pixels > 0 even

Class 1:
1 <= L2 norm <= 2

Class 2:
Everything else

Class 1:
Three modes

Class 2:
Everything else

18

So far: Defined a (linear) <u>score function</u> f(x,W) = Wx + b

Example class
scores for 3
images for
some W:

How can we tell
whether this W
is good or bad?

image by Nikita is licensed under CC-BY 2.0
image is CC0 1.0 public domain
g image is in the public domain

	Image 1	Image 2	Image 3
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

19