ntic master
revolucionamos la comunicación

UNIVERSIDAD COMPLUTENSE DE MADRID

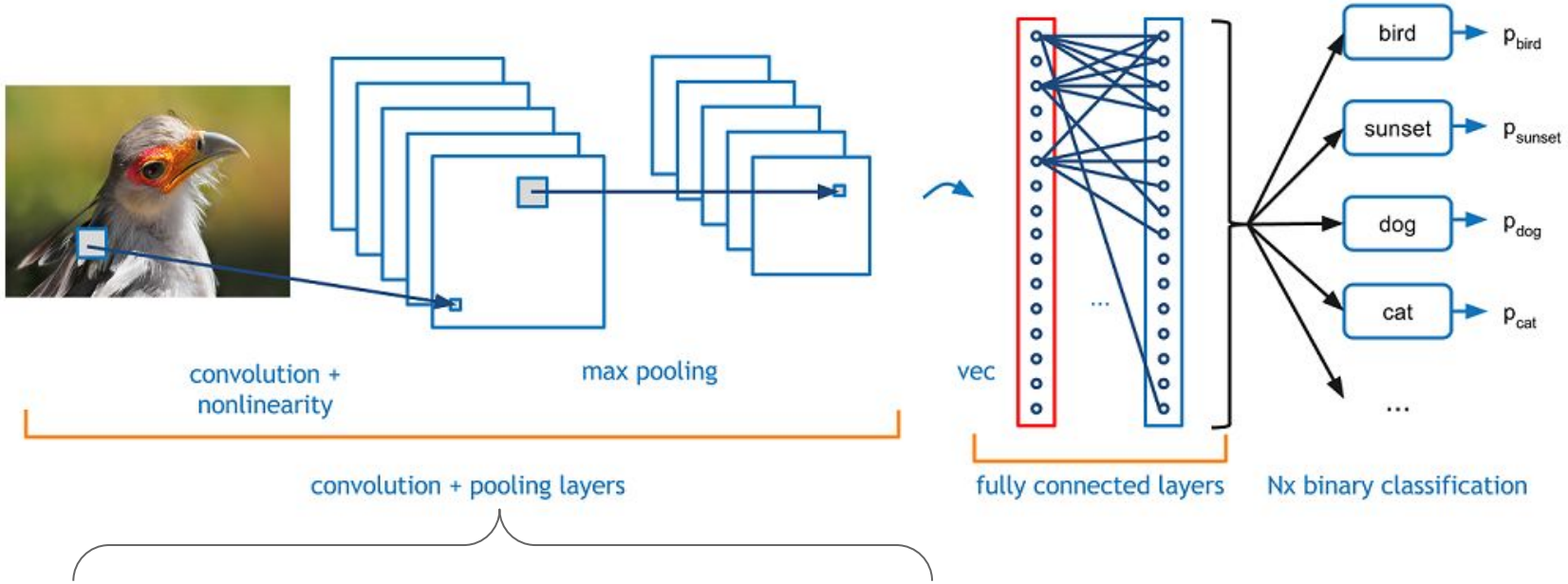
Deep Learning

Alberto Ezpondaburu

3

Convolutional Neural Networks (CNN) Use Cases

Feature Learning for Many Applications



Feature Learning for:
Image Classification, Object Detection, Image Segmentation, …

Image Classification Advances

Multi-class classification

Dog

Cat

Multi-Label classification

Dog, cat

Image Classification Advances

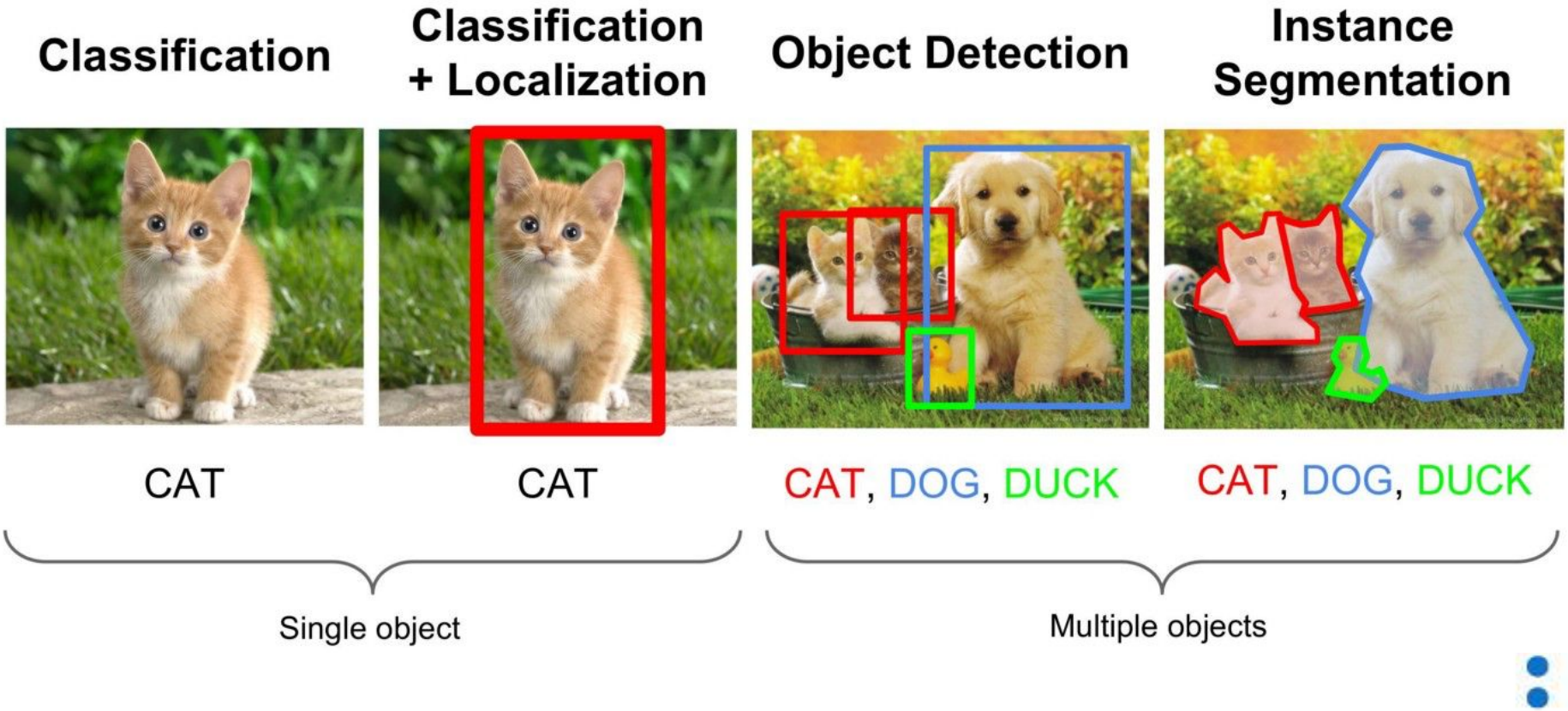
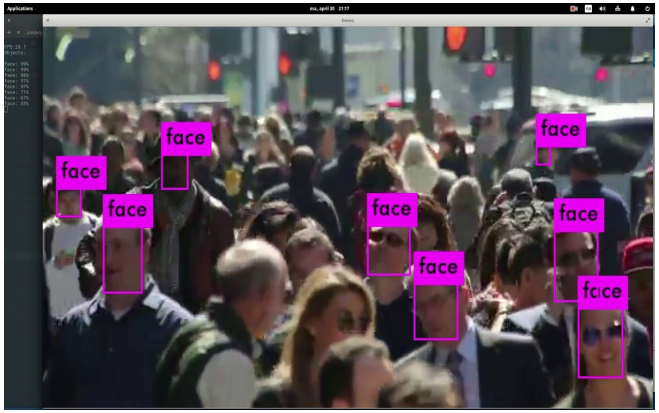
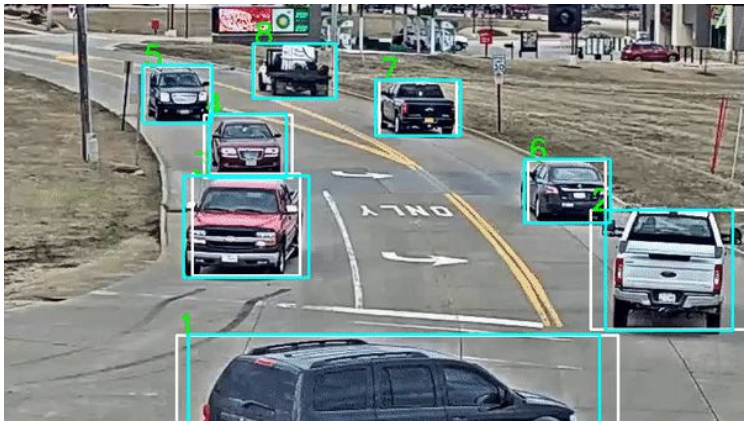
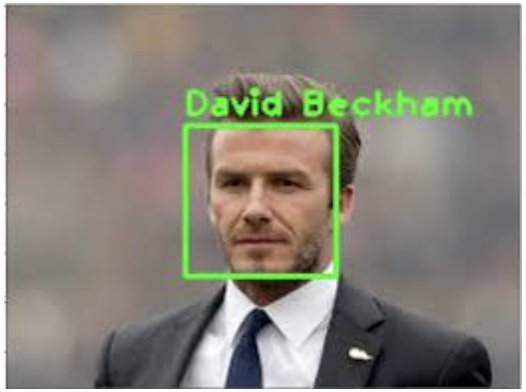


Image Classification Advances: Applications







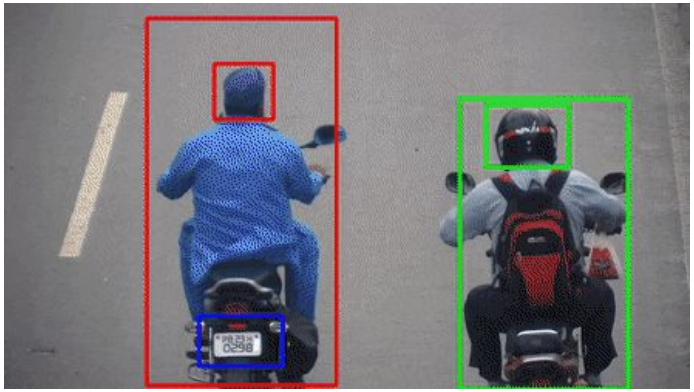
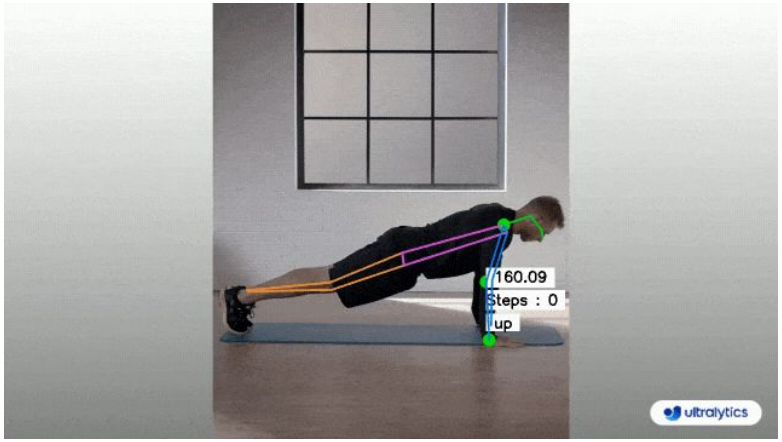
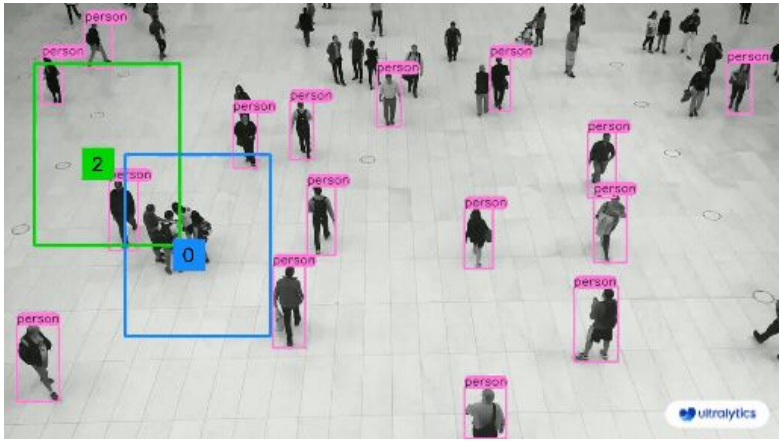


Image Classification Advances: Applications

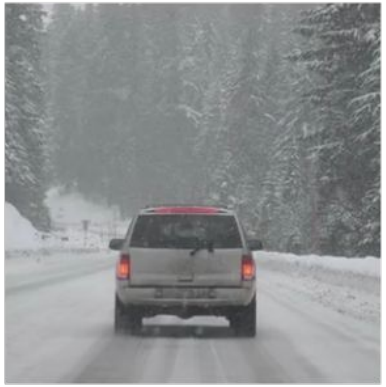




https://docs.ultralytics.com/

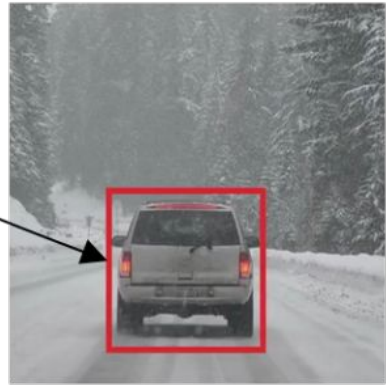
Object Localization I

Image classification



"Car"

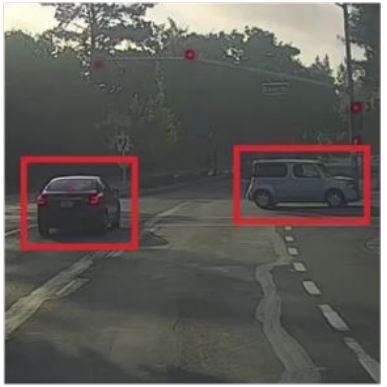
Classification with localization



"Car"

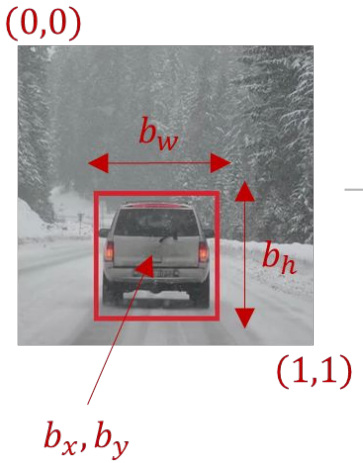
One object

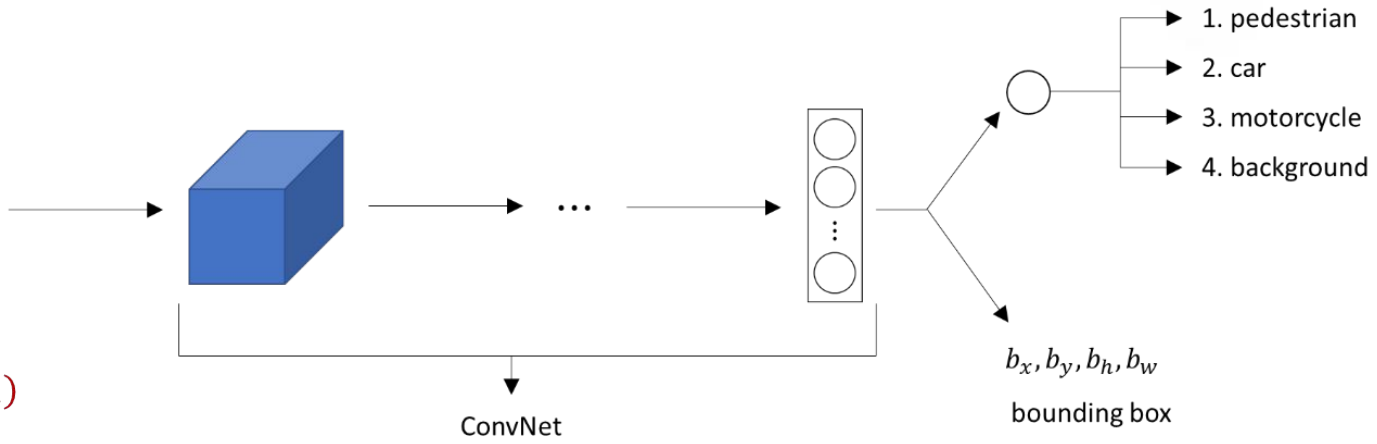
Detection



Multiple objects

Object Localization II





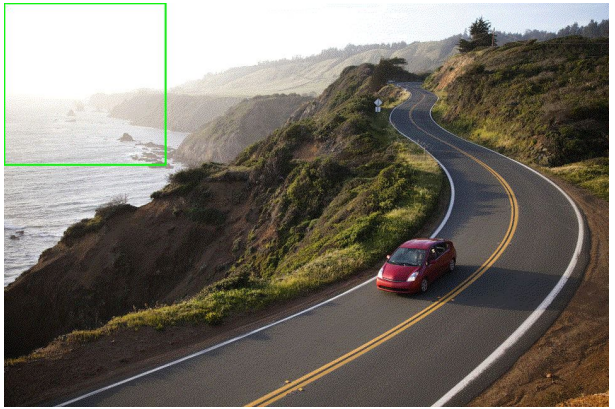
Target: $y = \begin{bmatrix} c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$

Sigmoid: Is there an object?

Bounding box, ReLu

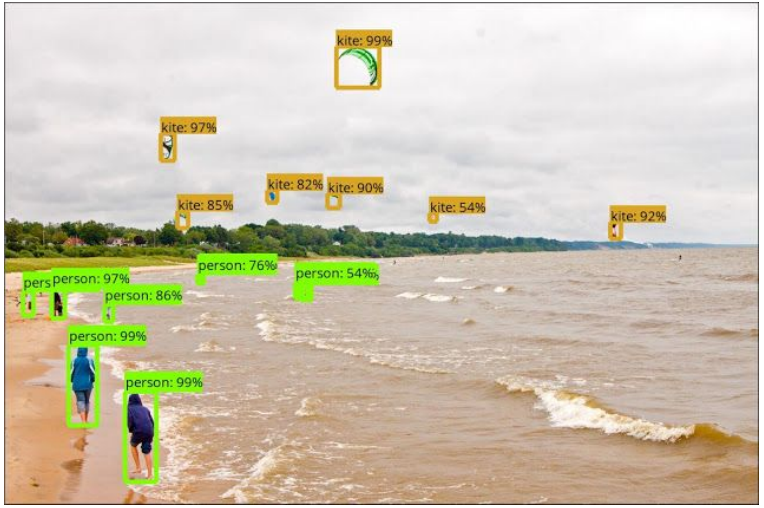
Classes of objects, softmax

Object Detection: Sliding Windows



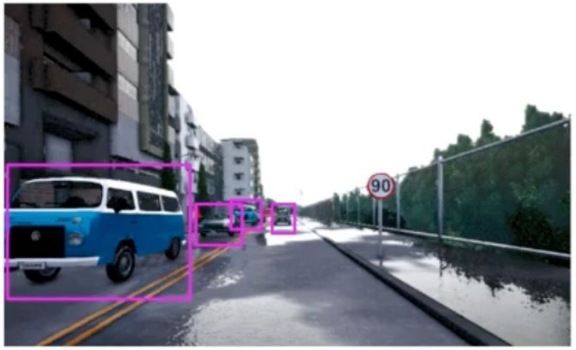
Object Detection: Popular Algorithms

- R-CNN
- Faster -RCNN
- YOLO
- SSD

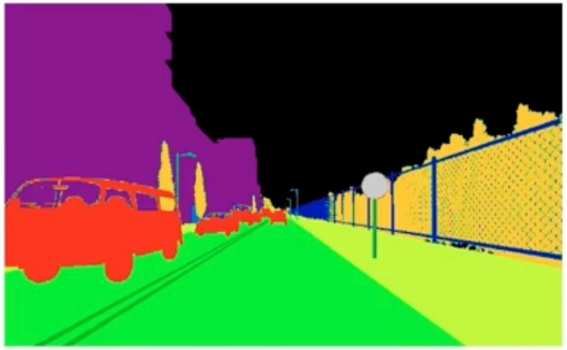


Semantic Segmentation vs Object Detection

Input image

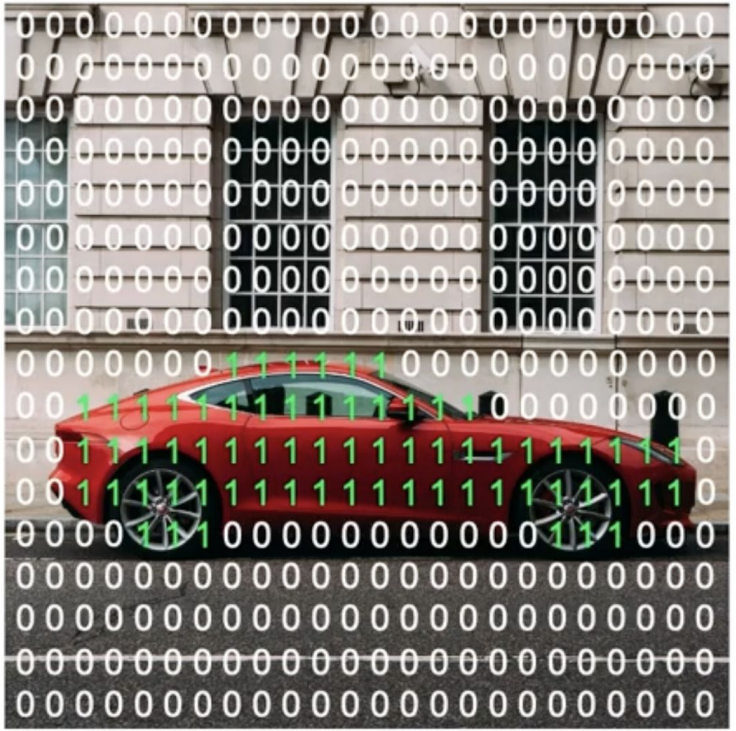


Object Detection



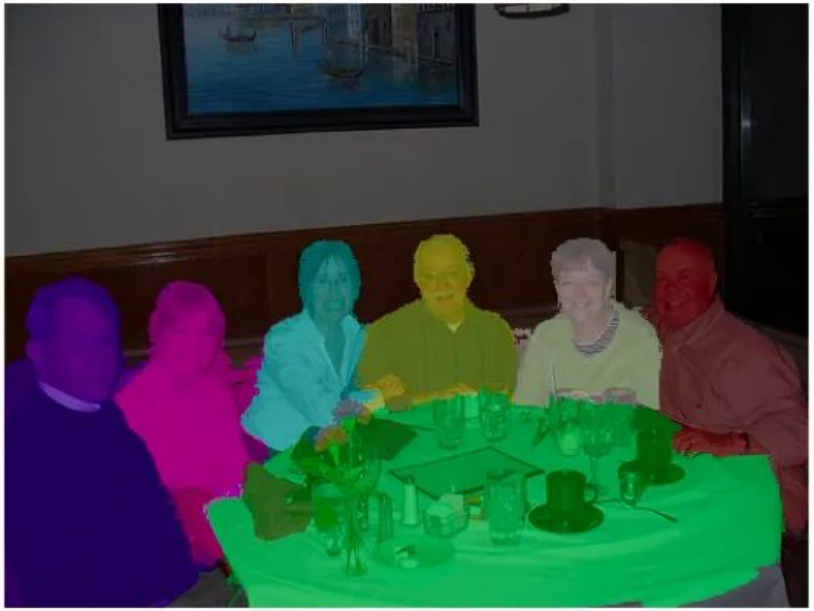
Semantic Segmentation

Semantic Segmentation: Per Pixel Class Labels



Semantic Segmentation vs Instance Segmentation

Semantic Segmentation



Instance Segmentation

SAM: Segment Anything Model

https://segment-anything.com/

Neural Style Transfer

content image | style image | generated image


(C)

+

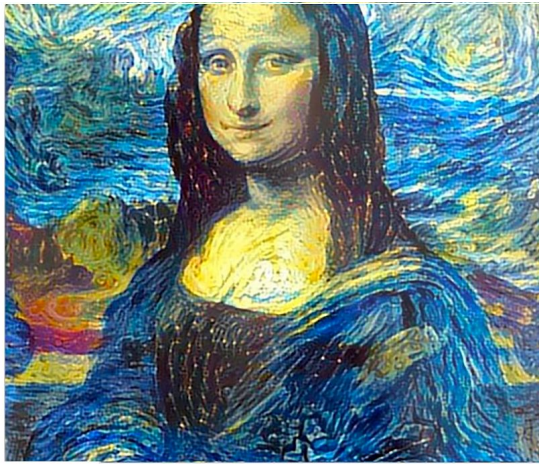

(S)

=


(G)

+

=



Paper : A Neural Algorithm of Artistic Style

Neural Style Transfer: Visualizing Deep Layer

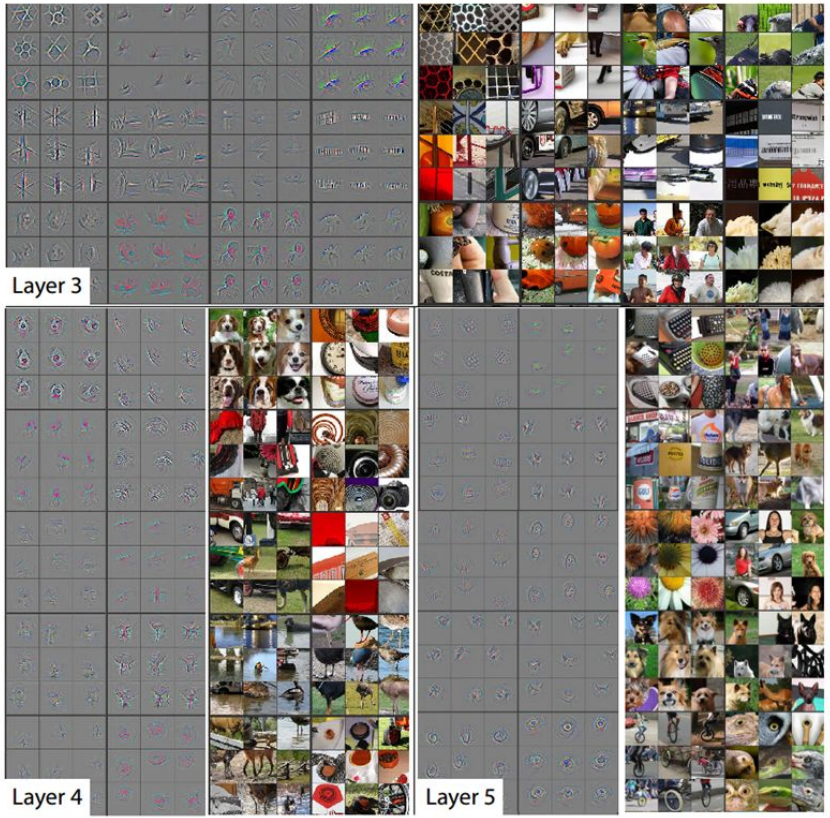
1. Use pre-trained model for extracting content and style (VGG-19).
2. Inputs: a pair of images, style and content.
   a. Extract style from image S
   b. Extract content from image C
3. Generate a new imagen with a training loop (gradient descend)
   a. Optimize Generated image G

(C)

(S)

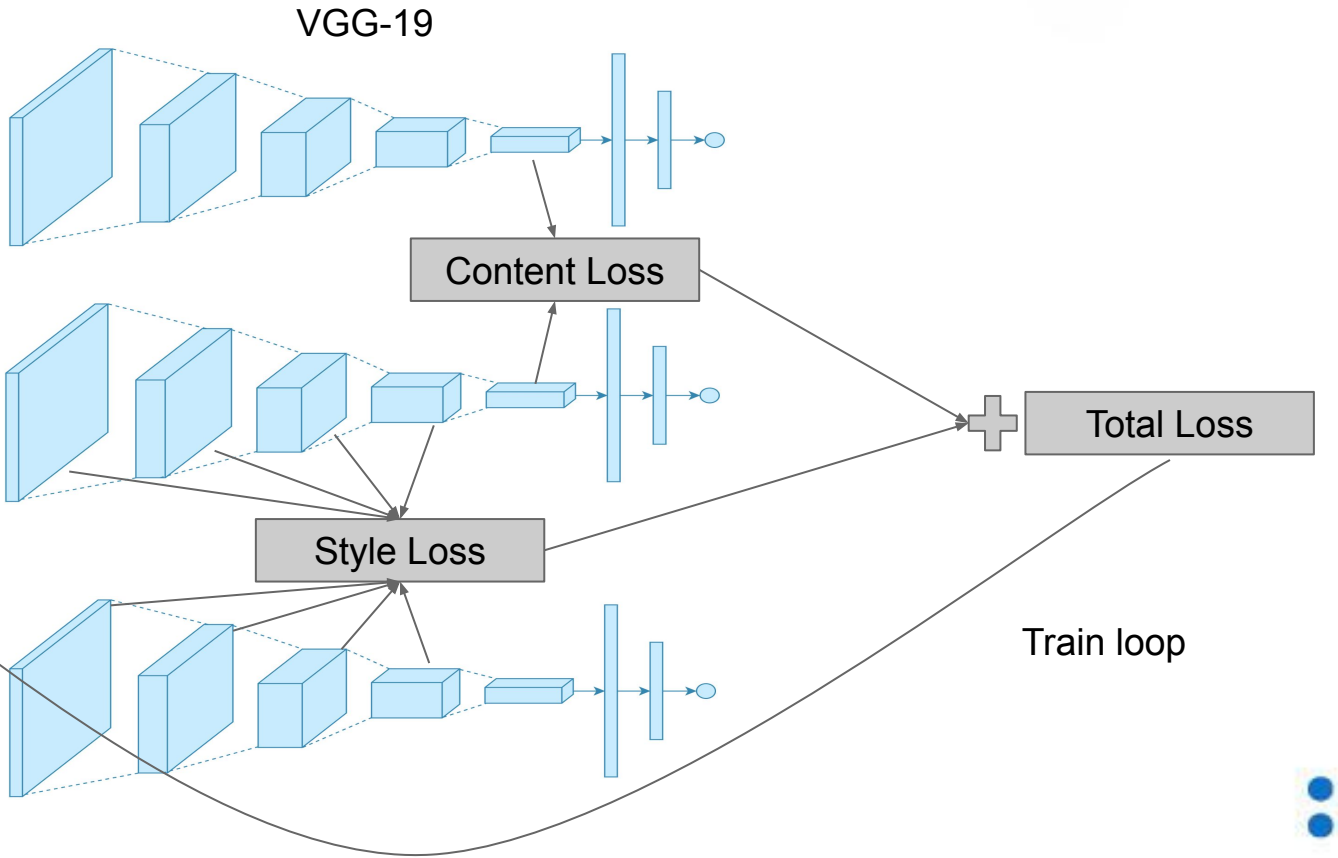
(G)

Neural Style Transfer: Visualizing Layers

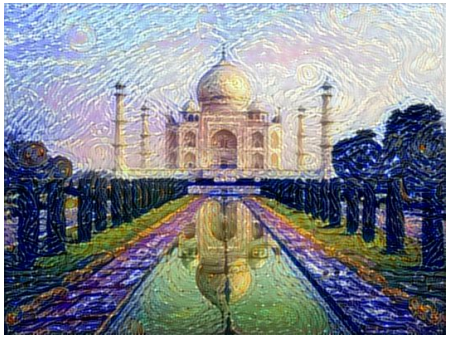


Visualizations of Layers 3, 4, and 5

Neural Style Transfer: Visualizing Deep Layer

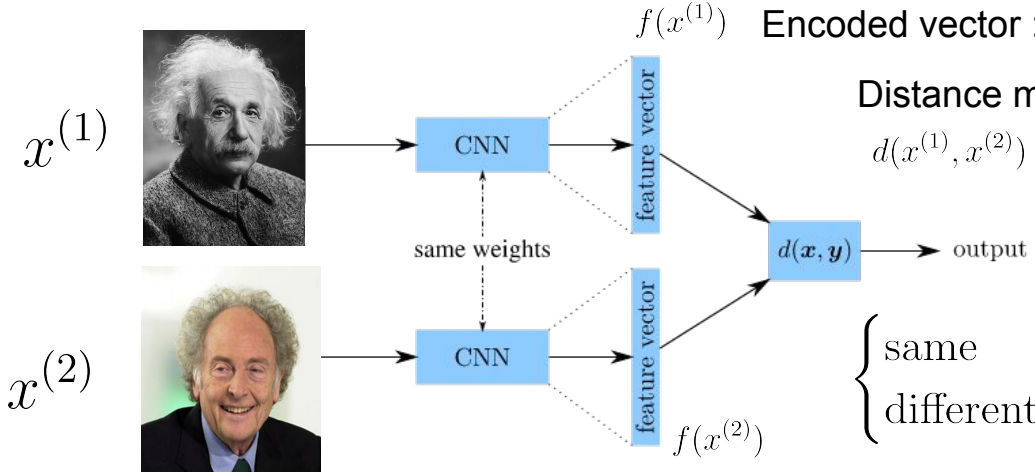


Neural Style Transfer



Face Recognition/Verification: Siamese Net

Verification: Input: Image and name => Correct/Wrong
Recognition: Input: Image => name (K possibilities) (much harder)
One-shot learning: Recognize a person given a single image. Small training set is not enough for training a CNN. Not need a classifier, use similarity function.

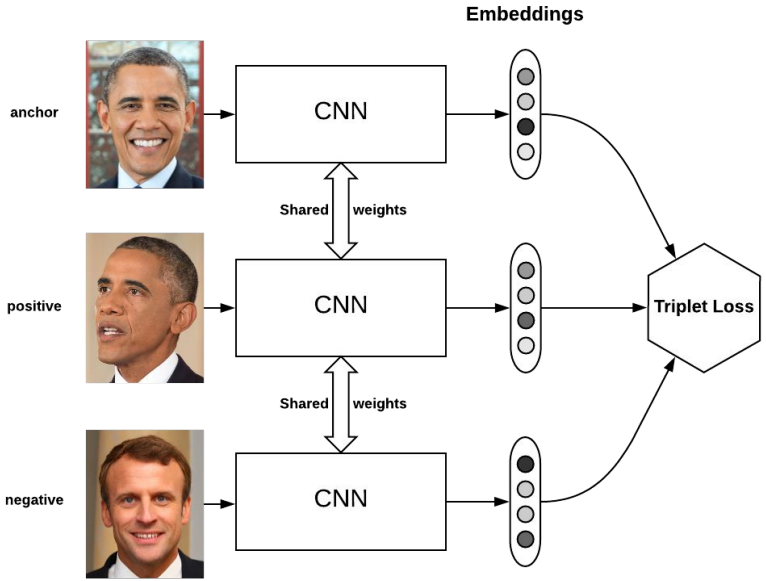


Distance measure between images

$$d(x^{(1)}, x^{(2)}) = ||f(x^{(1)}) - f(x^{(2)})||^2$$

$$\begin{cases} \text{same} & \text{if } d(x^{(1)}, x^{(2)}) \leq T \\ \text{different} & \text{if } d(x^{(1)}, x^{(2)}) > T \end{cases}$$

Face Recognition/Verification: Triplet Loss



- Train triplets: Anchor, Negative and Positive.
- We want the distance between anchor and positive be lower than the distance between the anchor and the negative image.

$$d(A, P) + \alpha < d(A, N)$$

- Train with more similars negative examples.

$$\sum_i \max\left(||f(A^{(i)}) - f(P^{(i)})||^2 - ||f(A^{(i)}) - f(N^{(i)})||^2 + \alpha, 0\right)$$

Siamese Net: Other Applications

Handwritten Signature

Voice check

How old are you?
What is your age?

Question duplicates

ntic master
revolucionamos la comunicación

UNIVERSIDAD COMPLUTENSE DE MADRID