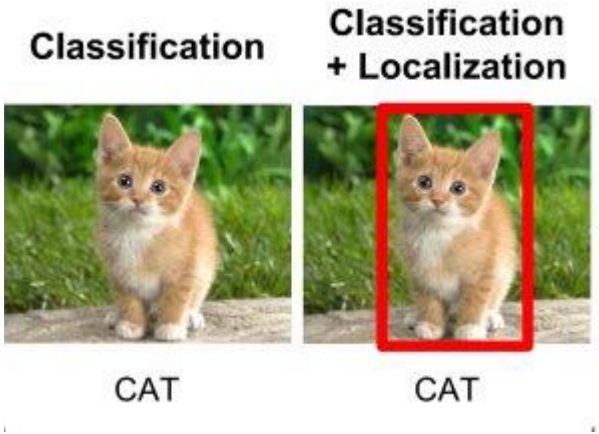
# TALLER DE DEEP LEARNING

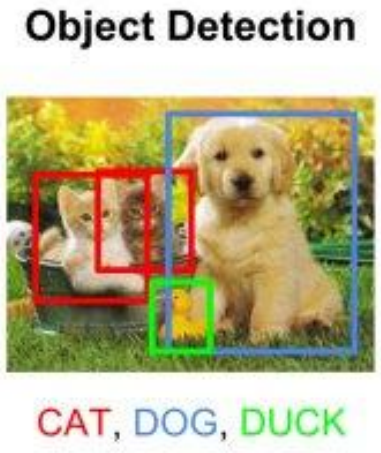
# Lectura 4:
# Conceptos básicos de detección de objetos

Dennis Núñez Fernández

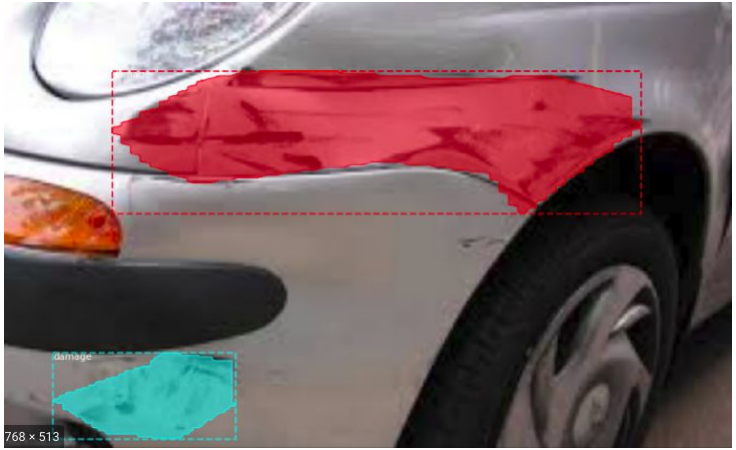
# Detección de Objetos



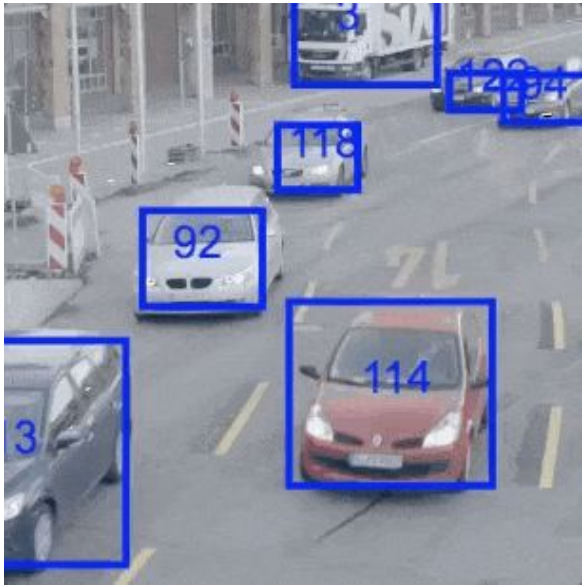
# Detección de Objetos



# Aplicaciones

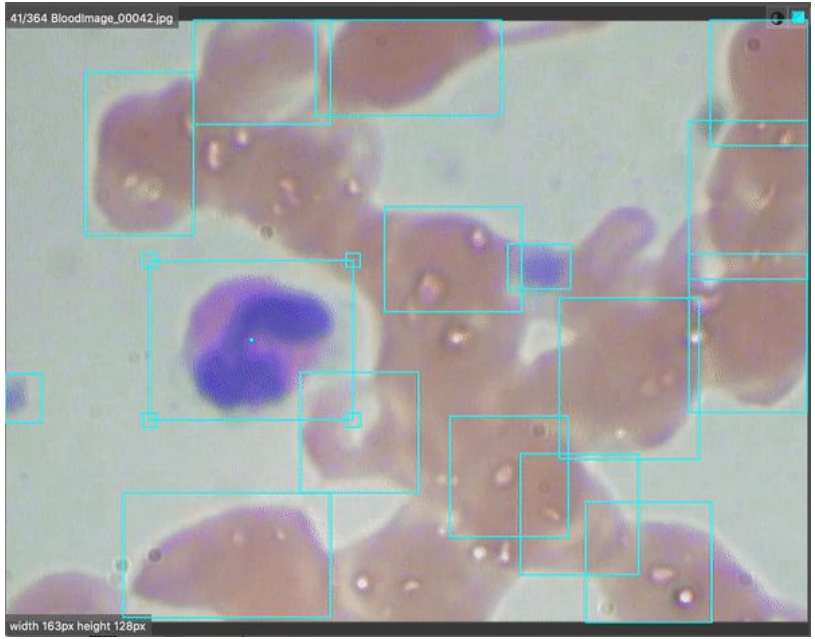
# Aplicaciones

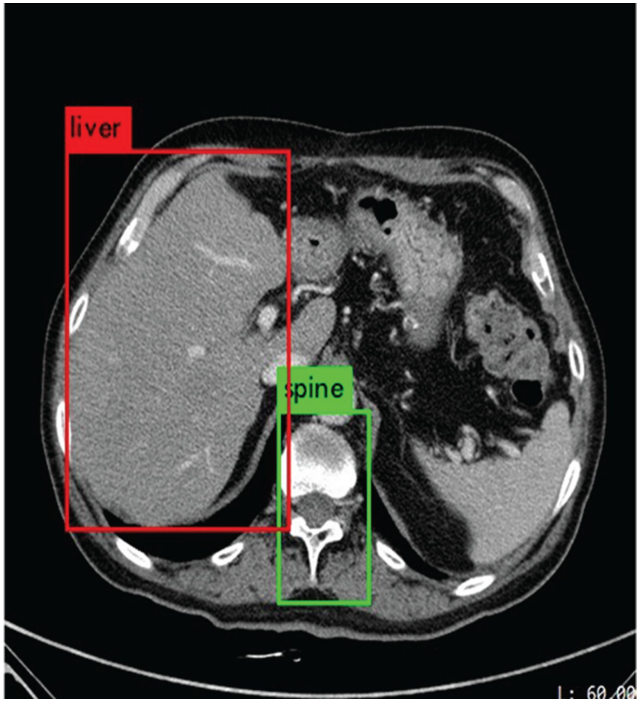



# Aplicaciones

# Aplicaciones

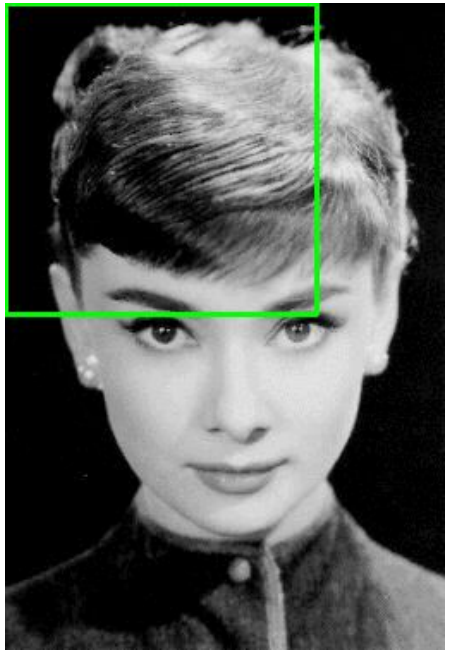

# Aplicaciones



# Detección de Objetos

Object Detection is modeled as a **classification problem** where we take windows of fixed sizes from input image at all the possible locations feed these patches to an image classifier.

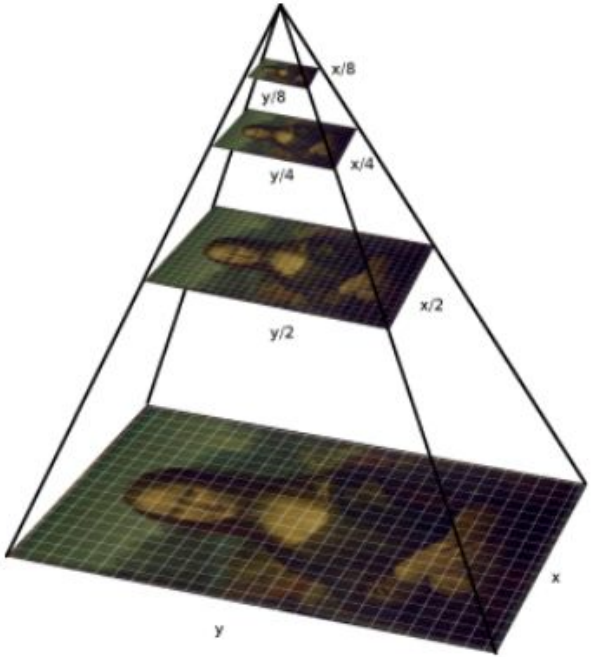


# Detección de Objetos

*Small sized object*

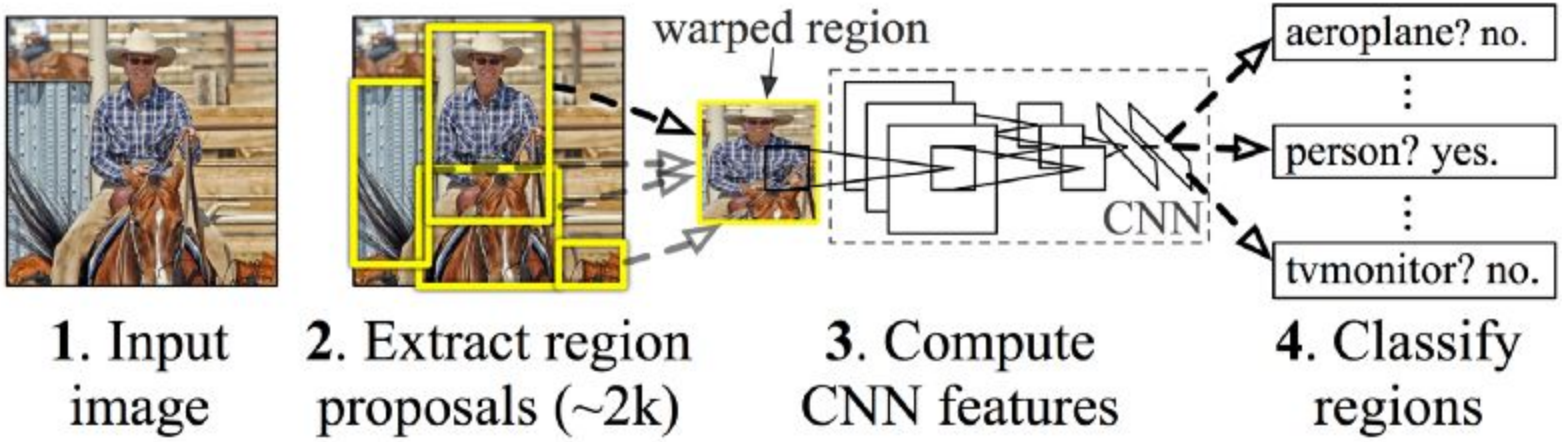
*Big sized object. What size do you choose for your sliding window detector?*

# Detección de Objetos



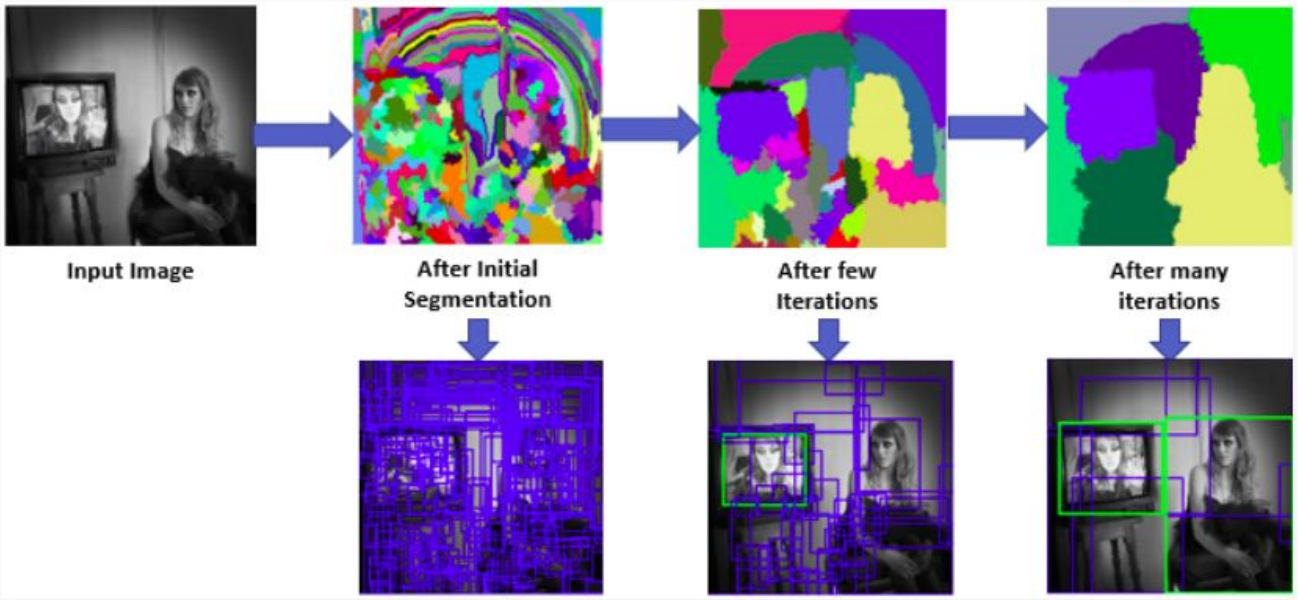
# Detección de Objetos

# R-CNN

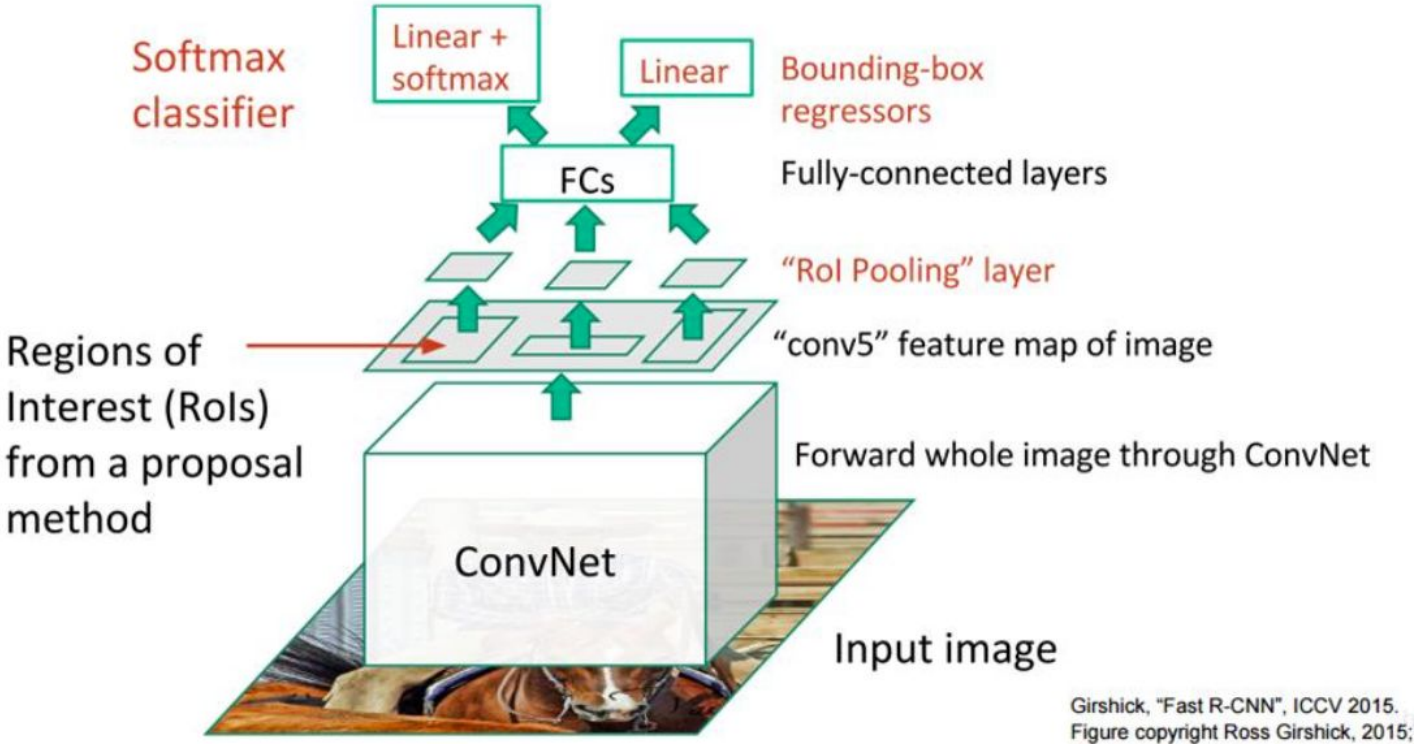

warped region
aeroplane? no.
person? yes.
tvmonitor? no.
CNN
1. Input image
2. Extract region proposals (~2k)
3. Compute CNN features
4. Classify regions

# R-CNN

Selective search algorithm:

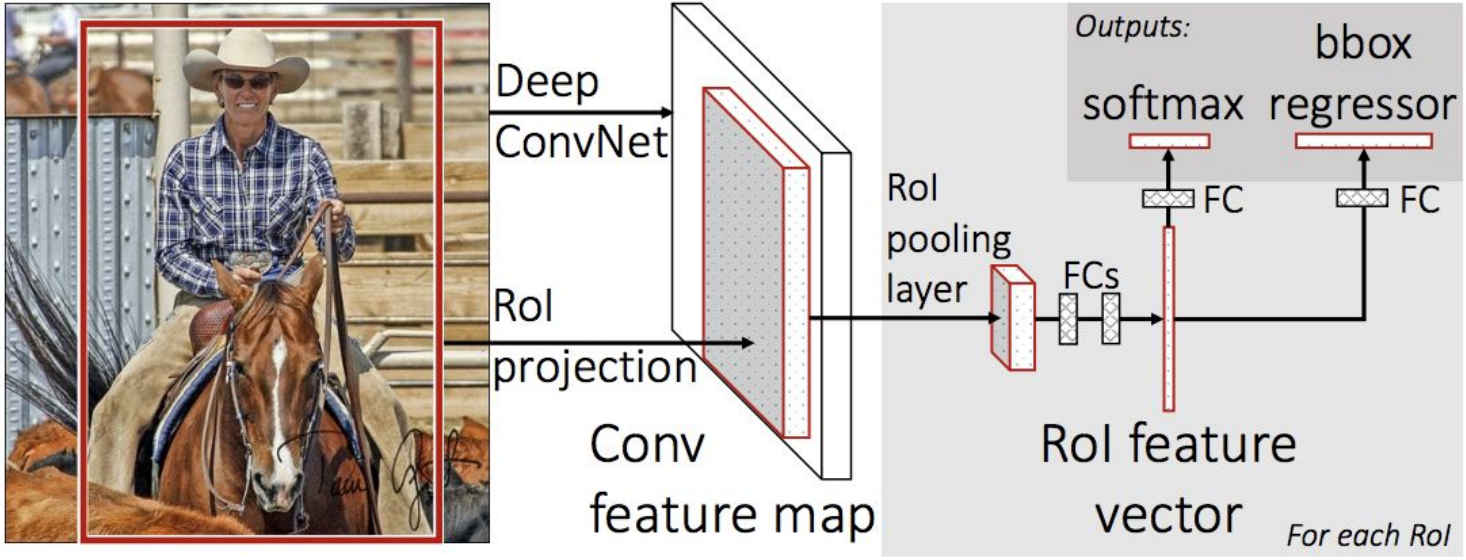


# Fast R-CNN

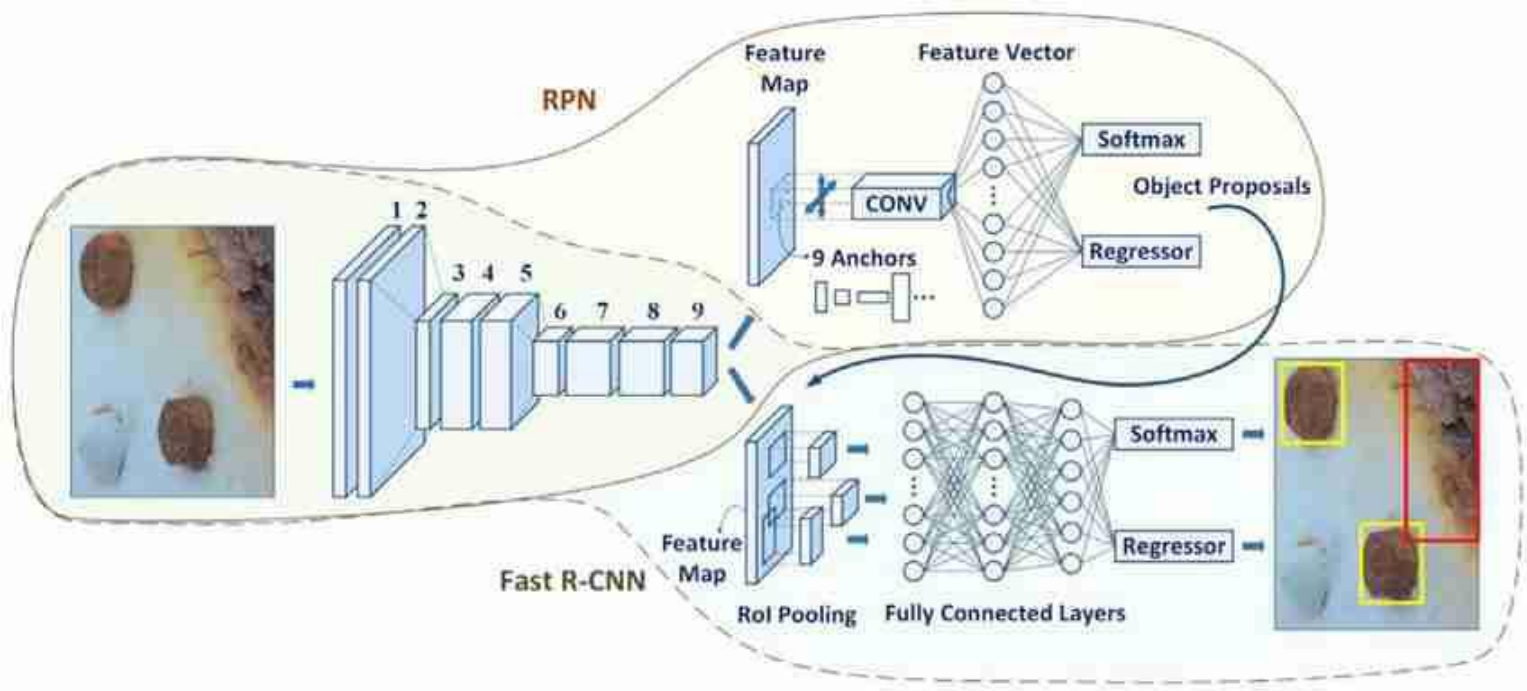



Girshick, "Fast R-CNN", ICCV 2015.
Figure copyright Ross Girshick, 2015;

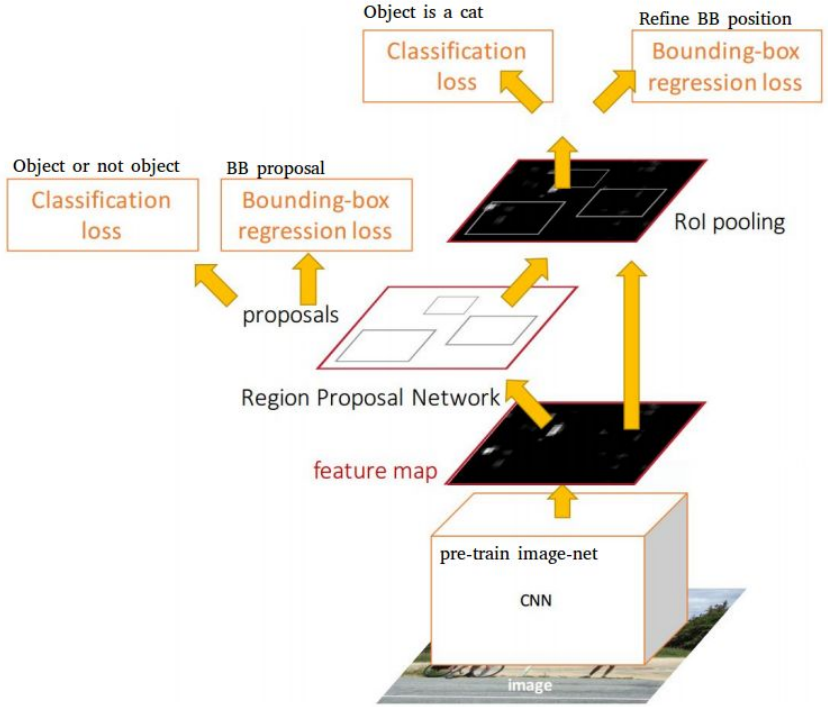
# Fast R-CNN


Deep
ConvNet
RoI
projection
Conv
feature map
RoI
pooling
layer
FCs
FC
FC
Outputs:
softmax
bbox
regressor
RoI feature
vector
For each RoI

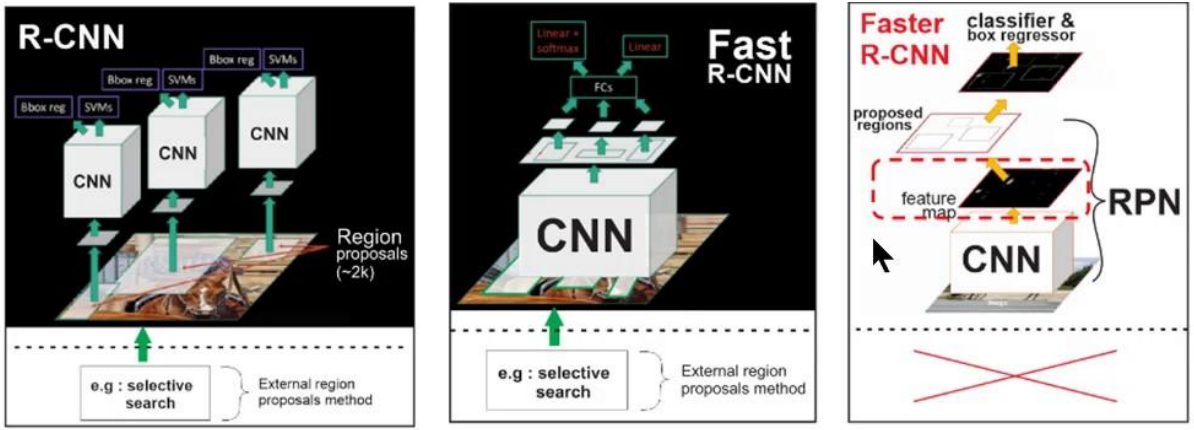
# Faster R-CNN


RPN
Feature Map
Feature Vector
Softmax
Object Proposals
CONV
Regressor
9 Anchors
1 2
3 4 5
6 7 8 9
Feature Map
Fast R-CNN
RoI Pooling
Fully Connected Layers
Softmax
Regressor

# Faster R-CNN


Object is a cat
Classification loss
Refine BB position
Bounding-box regression loss
Object or not object
Classification loss
BB proposal
Bounding-box regression loss
RoI pooling
proposals
Region Proposal Network
feature map
pre-train image-net
CNN
image

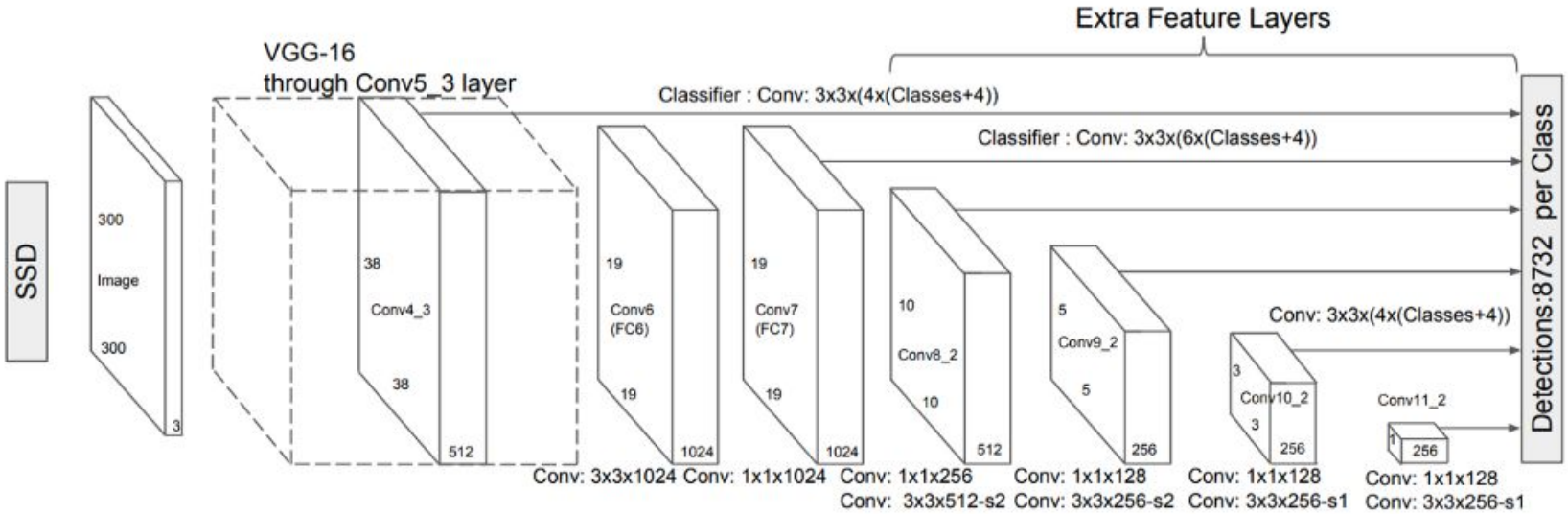
# Comparison



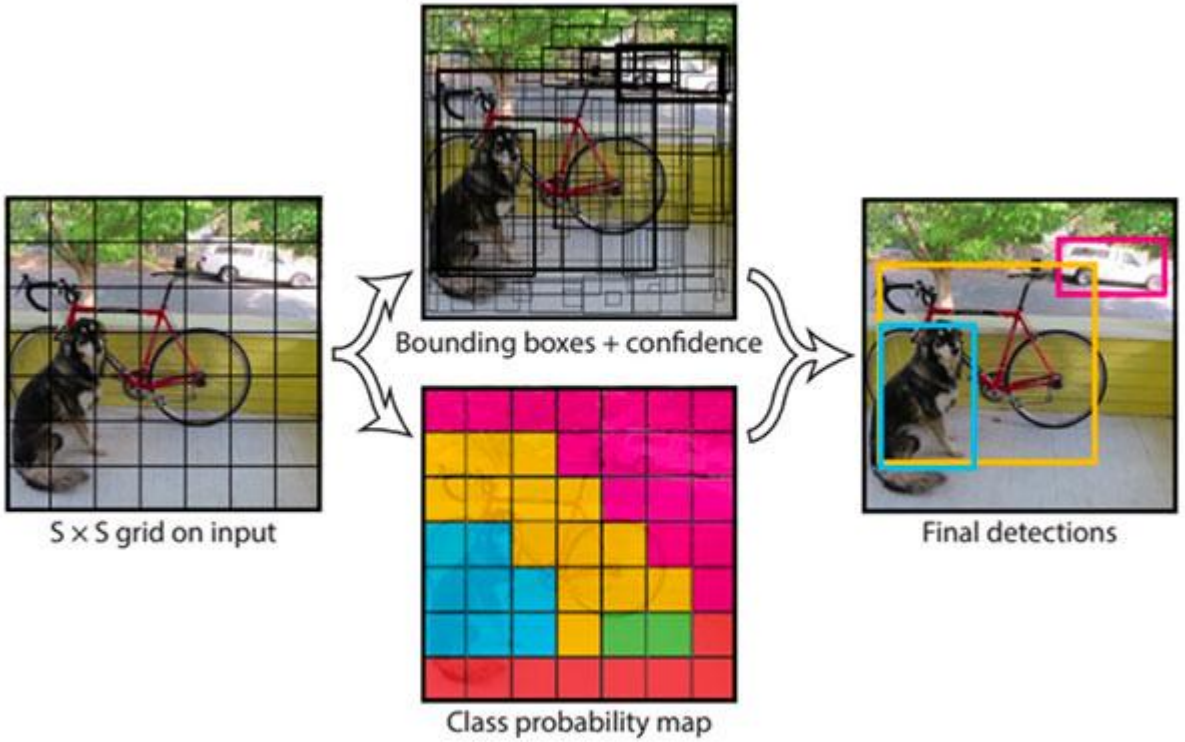

| | R-CNN | Fast R-CNN | Faster R-CNN |
|---|---|---|---|
| Test time per image | 50 seconds | 2 seconds | 0.2 seconds |
| Speed-up | 1x | 25x | 250x |
| mAP (VOC 2007) | 66.0% | 66.9% | 66.9% |

* Standford lecture notes on CNN by Fei Fei Li and Andrej Karpathy

# SSD (Single Shot Detector)



# YOLO (You Only Look Once)



# Metrics

**‘Intersection over Union’ (IoU or overlap)**


$$\text{Score} = \frac{\text{Area of overlap}}{\text{Area of union}}$$

# Detección de Objetos

# Tools: VGG annotator