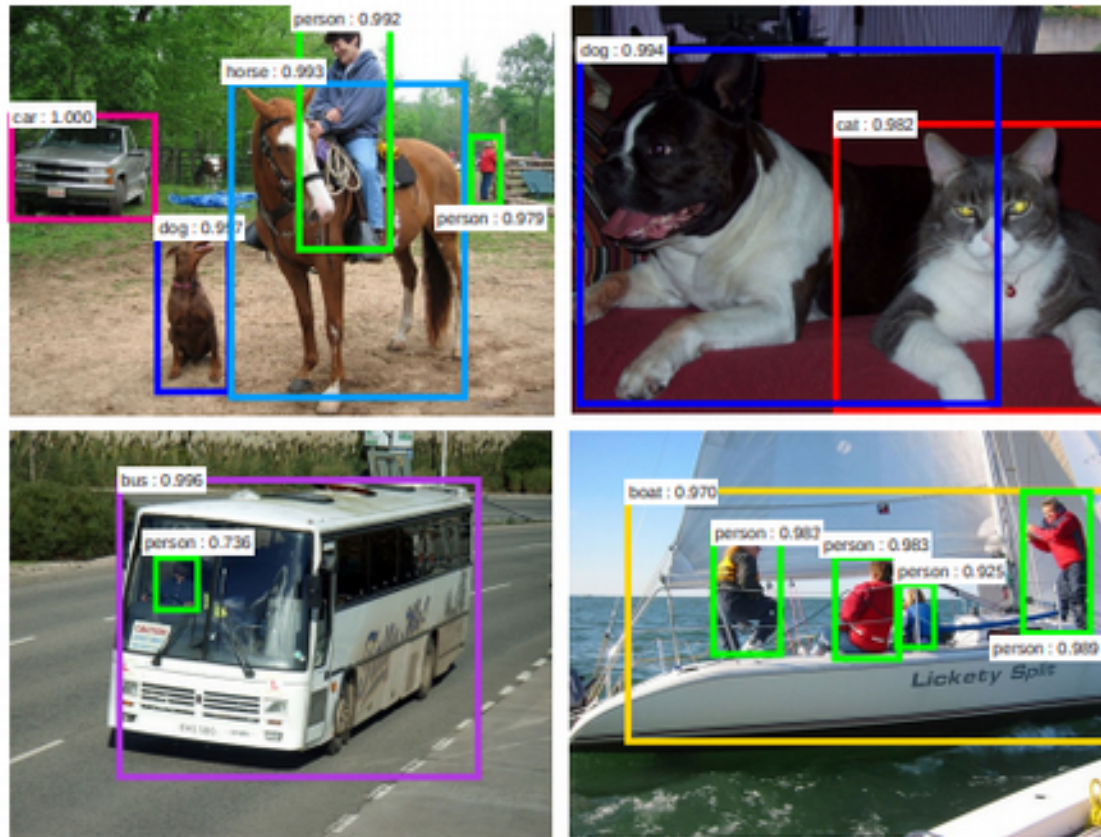


DETECCIÓN de OBJETOS

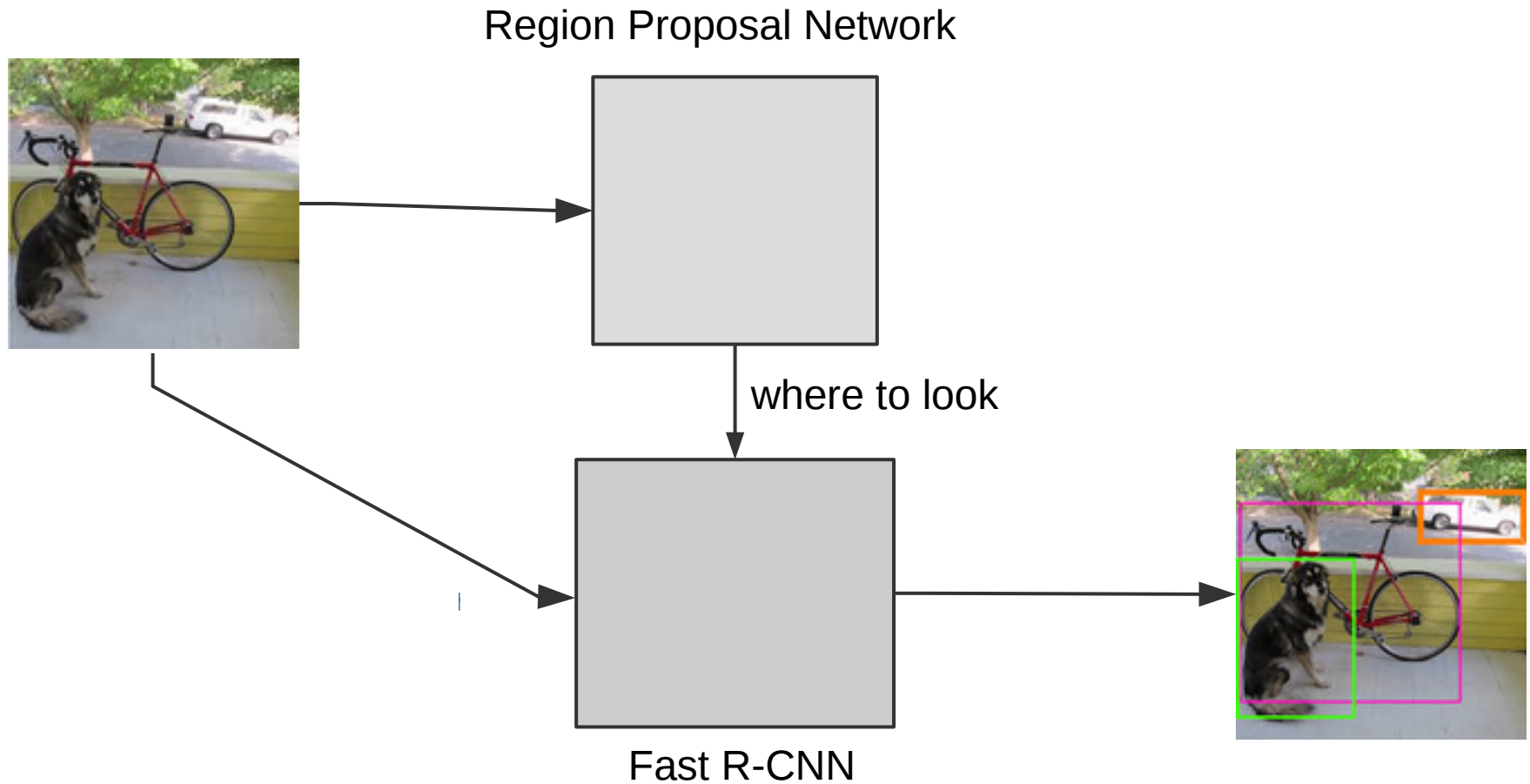
Parte I: Faster RCNN

José M. Saavedra R.

Detección de Objetos



Faster R-CNN



Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks

Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun

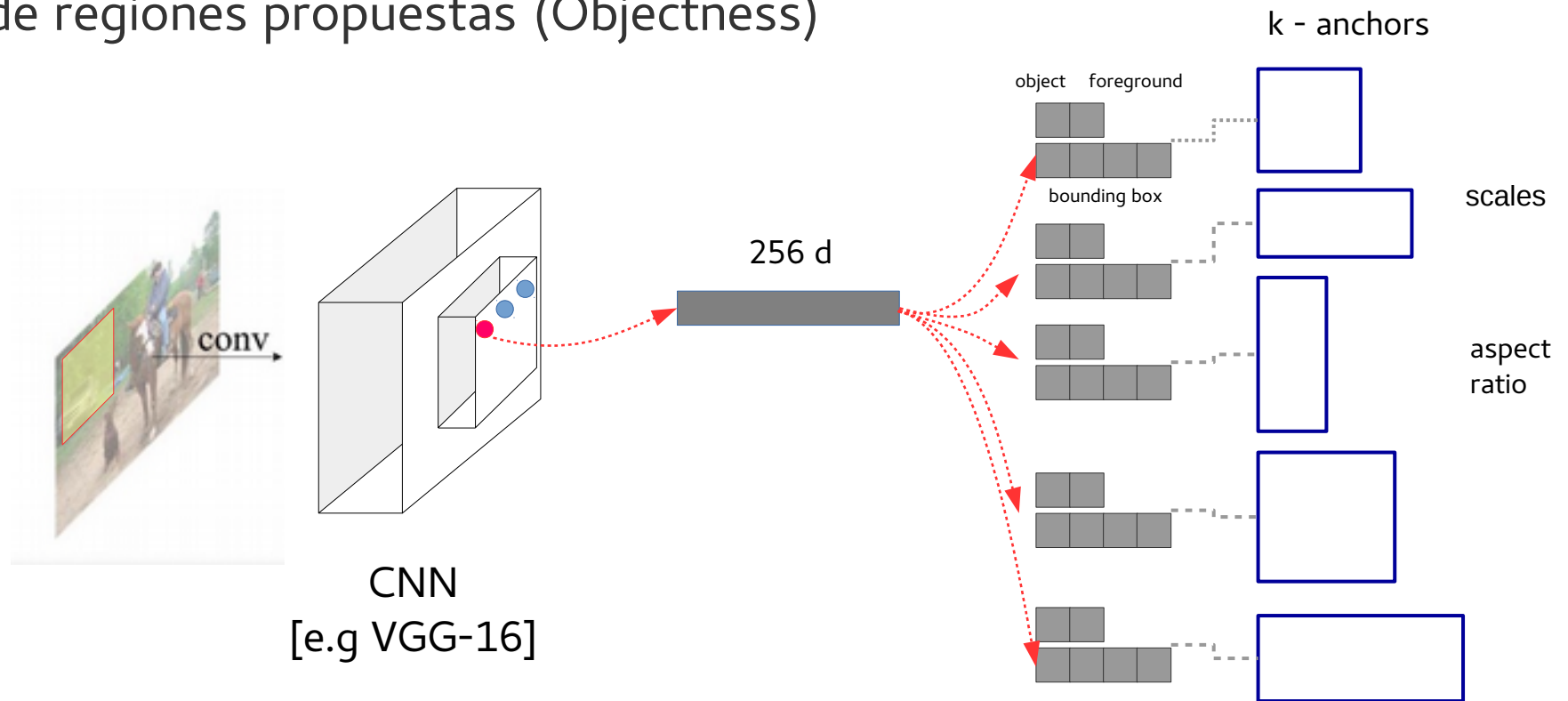
<https://arxiv.org/pdf/1506.01497.pdf>

CC6204 José M. Saavedra

Faster R-CNN

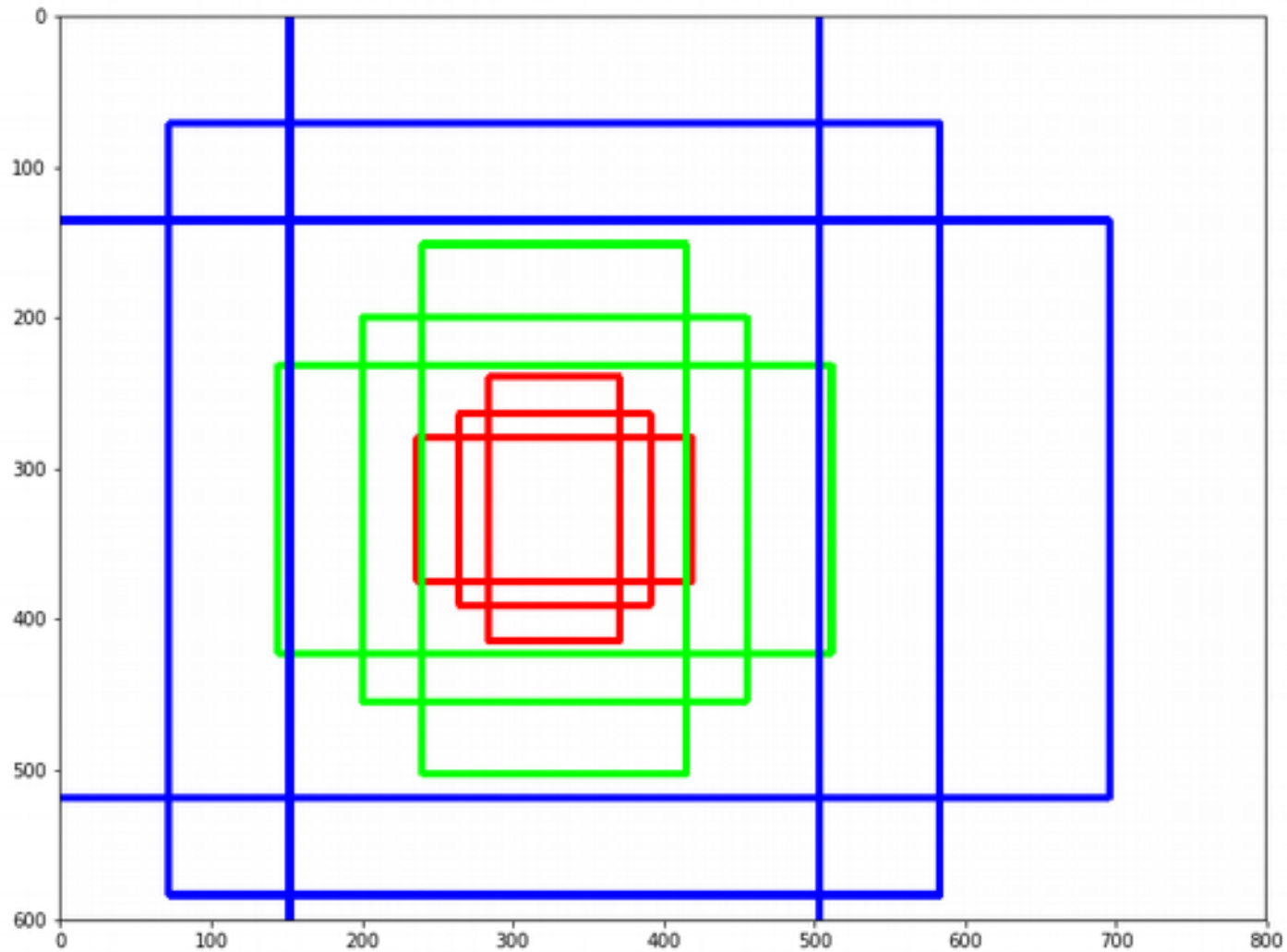
Region Proposal Network

Una RPN recibe una imagen de cualquier tamaño y retorna un conjunto de regiones propuestas (Objectness)



FASTER R-CNN

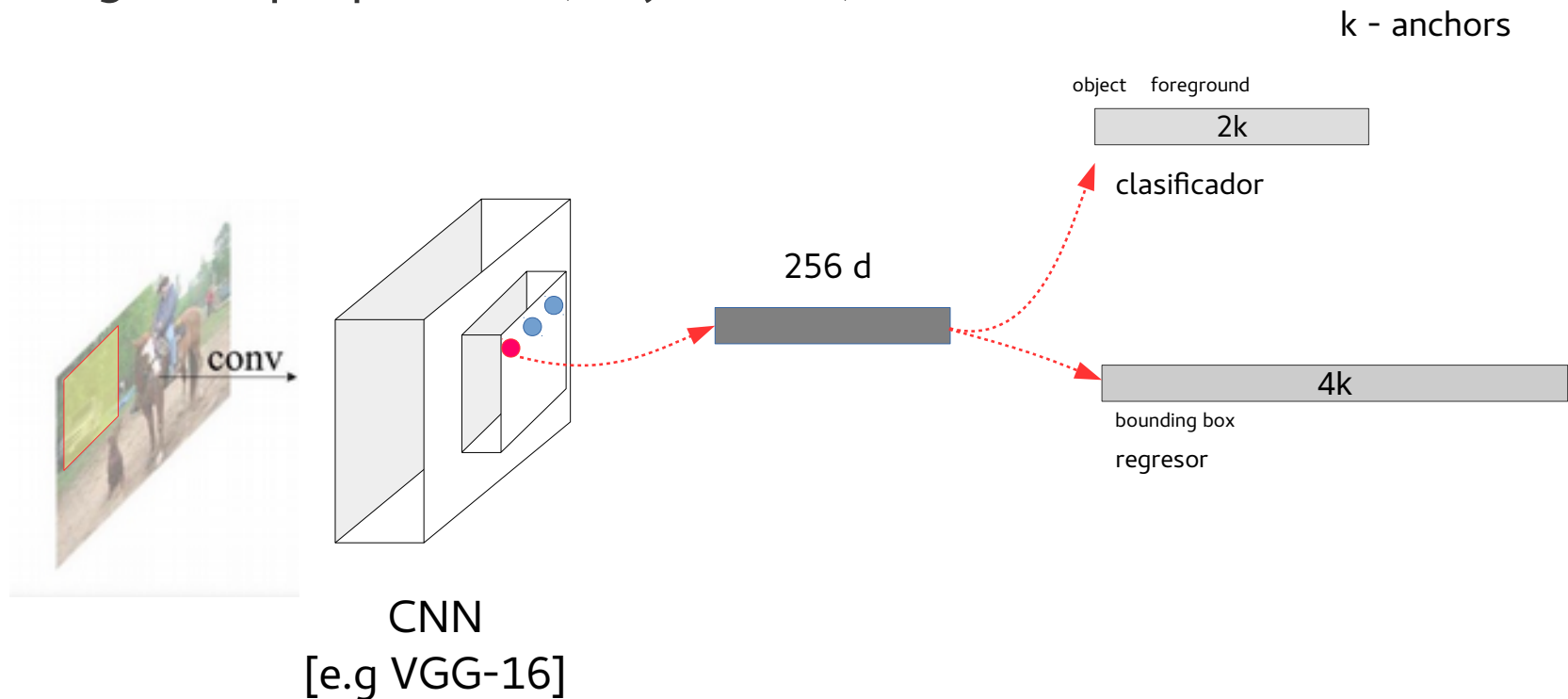
ANCHORS



Faster R-CNN

Region Proposal Network

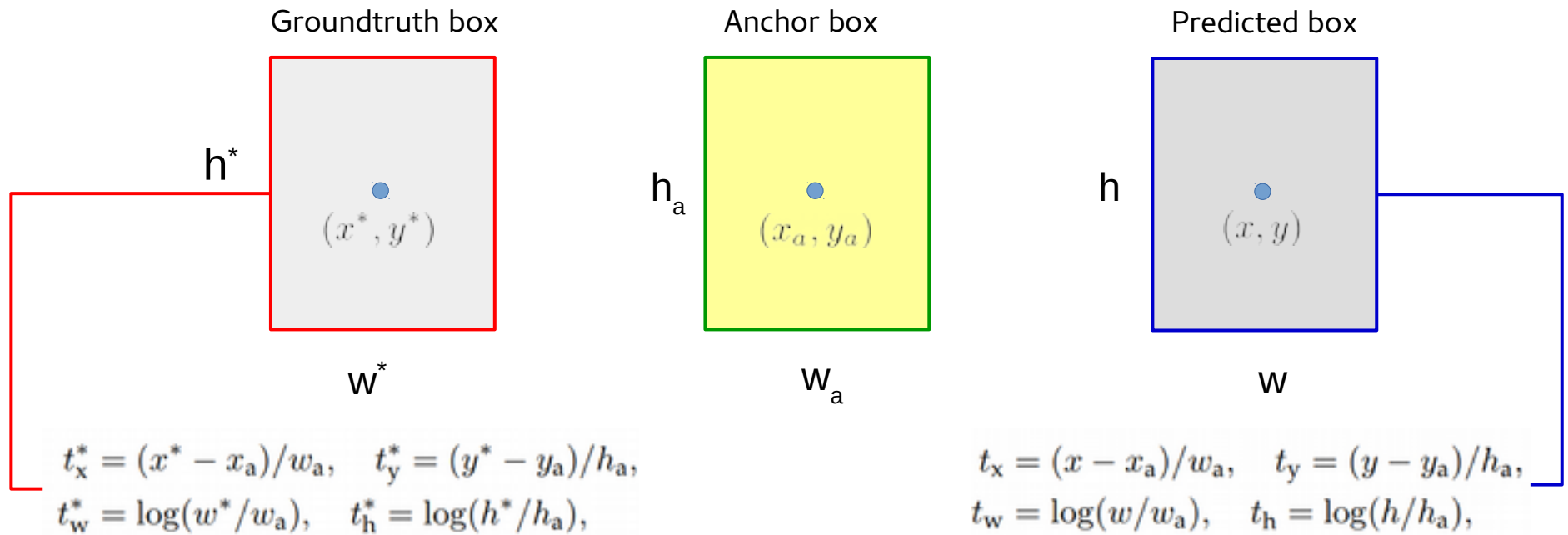
Una RPN recibe una imagen de cualquier tamaño y retorna un conjunto de regiones propuestas (Objectness)



Faster R-CNN

Region Proposal Network

ANCHOR



t_x	t_y	t_w	t_h
-------	-------	-------	-------

Faster R-CNN

Region Proposal Network

Loss

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) \\ + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*).$$

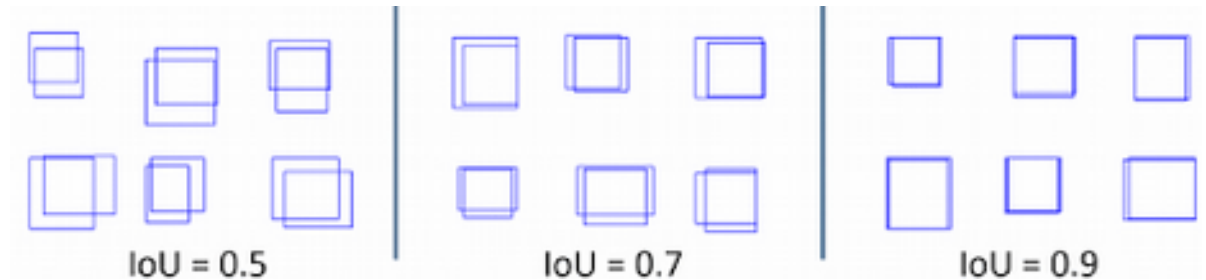
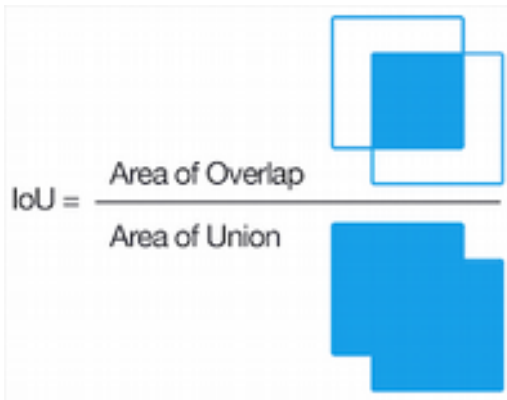
$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*)$$

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise,} \end{cases}$$

Faster R-CNN

Region Proposal Network

IoU: Intersection over Union

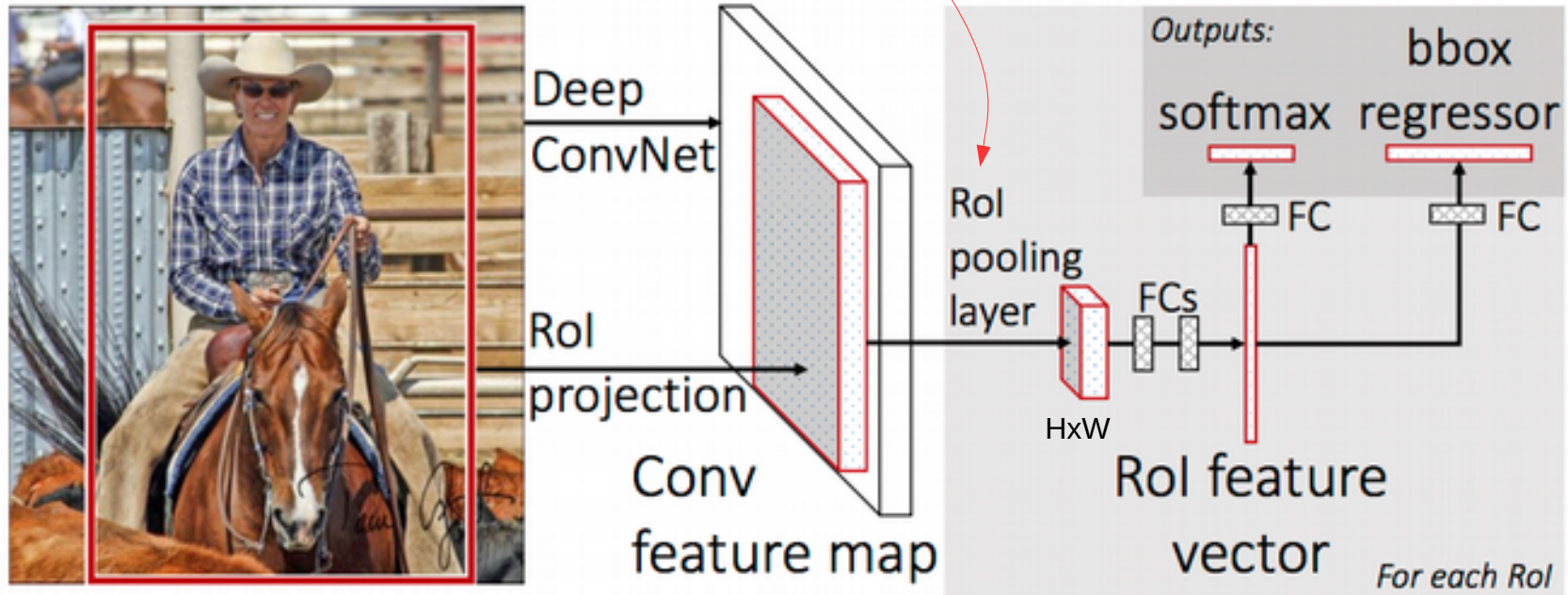


$\text{IoU} < 0.3 \rightarrow \text{background}$

Faster R-CNN

Fast R-CNN

Max-pooling, con vecindad variable



Fast R-CNN

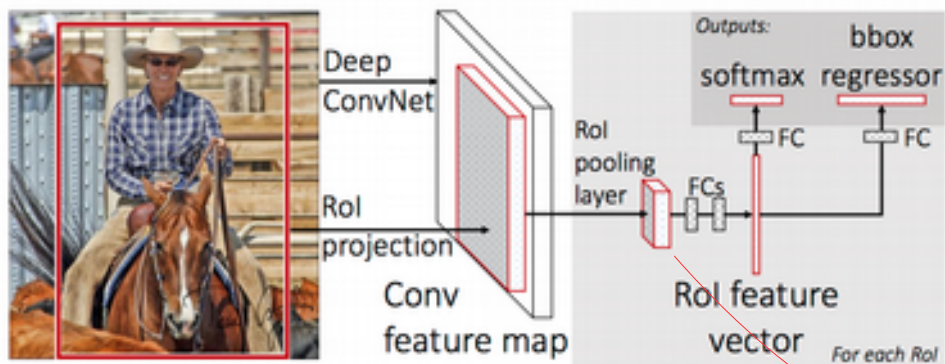
Ross Girshick
Microsoft Research
rbg@microsoft.com

CC6204 José M. Saavedra

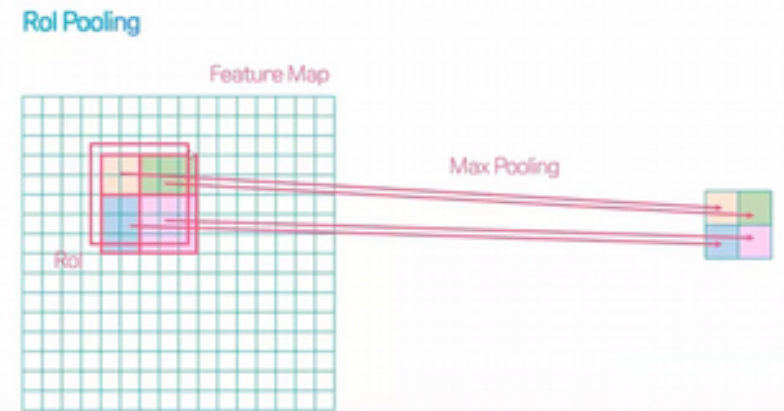
<https://arxiv.org/pdf/1504.08083.pdf>

Faster R-CNN

Fast R-CNN



Dividir ROI en HxW celdas



Fast R-CNN

Ross Girshick
Microsoft Research
rbg@microsoft.com

Clasificador: $K+1$ $K = \# \text{ clases}$
Regresor: $4K$

CC6204 José M. Saavedra

Faster R-CNN

Fast R-CNN

Con respecto a bbox, se procede en forma similar que para los achors. Siendo el punto de referencia el especificado por la región propuesta.

Región Propuesta

$$P = \{P_x, P_y, P_w, P_h\}$$

$$t_x^* = (G_x - P_x)/P_w$$

Región Groundtruth

$$t_y^* = (G_y - P_y)/P_y$$

$$G = \{G_x, G_y, G_w, G_h\}$$

$$t_w^* = \log(G_w/P_w)$$

$$t_h^* = \log(G_h/P_h)$$

Fast R-CNN

Ross Girshick
Microsoft Research
rbg@microsoft.com

CC6204 José M. Saavedra

Faster R-CNN

Fast R-CNN

Con respecto a bbox, se procede en forma similar que para los achors. Siendo el punto de referencia el especificado por la región propuesta.

Región Propuesta

$$P = \{P_x, P_y, P_w, P_h\}$$

$$t_x = (x - P_x)/P_w$$

$$t_y = (y - P_y)/P_h$$

Predicción ?

$$t_w = \log(w/P_w)$$

x = ? y = ? w = ? h = ?

$$t_h = \log(h/P_h)$$

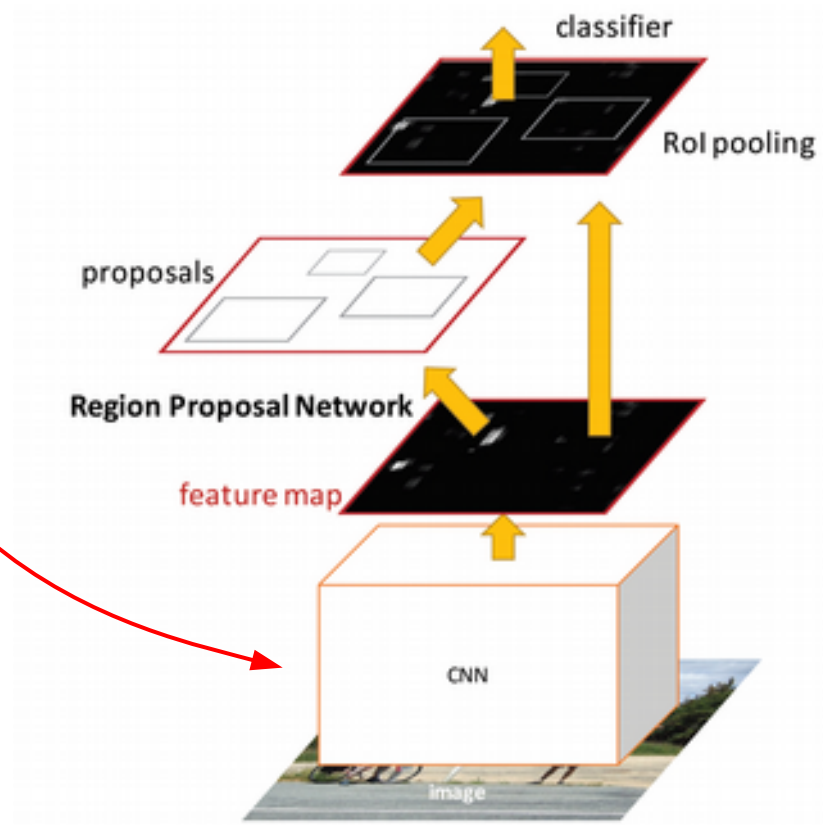
Fast R-CNN

Ross Girshick
Microsoft Research
rbg@microsoft.com

CC6204 José M. Saavedra

Faster R-CNN

Sharing Deep Features



Fast R-CNN

Ross Girshick
Microsoft Research
rbg@microsoft.com

<https://arxiv.org/pdf/1504.08083.pdf>

CC6204 José M. Saavedra