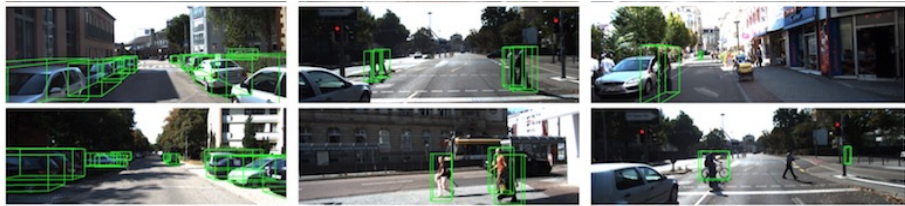


Una revisión sobre redes convolucionales para localización de objetos en imágenes

Julio Weissman

Licenciatura en Ciencias de la Computación
Universidad de Sonora

12 de abril de 2018



- ① ¿Qué es la localización de objetos en imágenes?
- ② Arquitectura RCNN
 - La arquitectura básica: RCNN
 - Su primer modificación: Fast RCNN
 - La que se emplea actualmente: Fastest RCNN
- ③ Arquitecturas alternativas
 - R-FCN
 - SSD
- ④ Conclusiones

Detección de objetos en imágenes

- 1 Redes convolucionales para tratamiento de imágenes
- 2 Reconocimiento de objetos en imágenes, la aplicación más exitosa de las CNN
- 3 Localización de objetos en imágenes es parecido a aplicar el reconocimiento en muchas subimágenes y luego fusionar la información

Redes convolucionales basadas en regiones (RCNN)

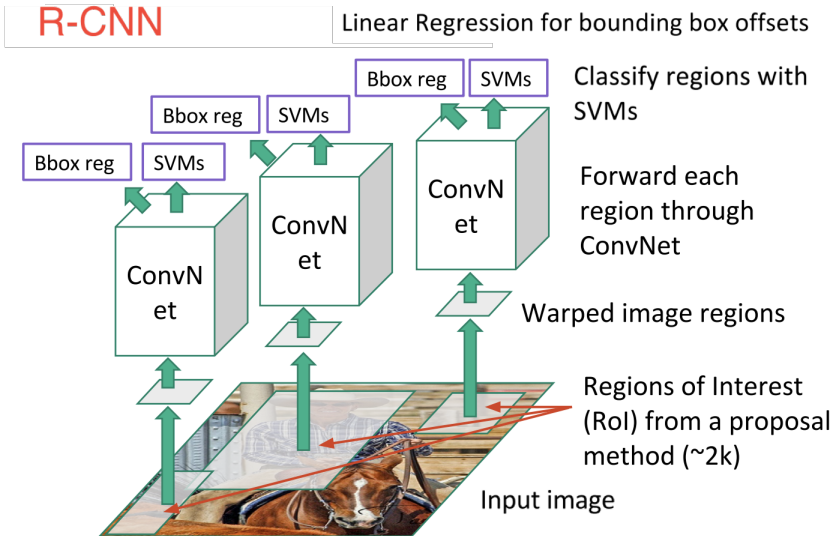
- 1 La primera red convolucional para detección de objetos
- 2 Es la base de todas las redes siguientes, pero solo se utiliza actualmente como método ilustrativo.
- 3 Muy lenta, implica el entrenamiento de muchos métodos

Redes convolucionales basadas en regiones (RCNN)

Algoritmo general

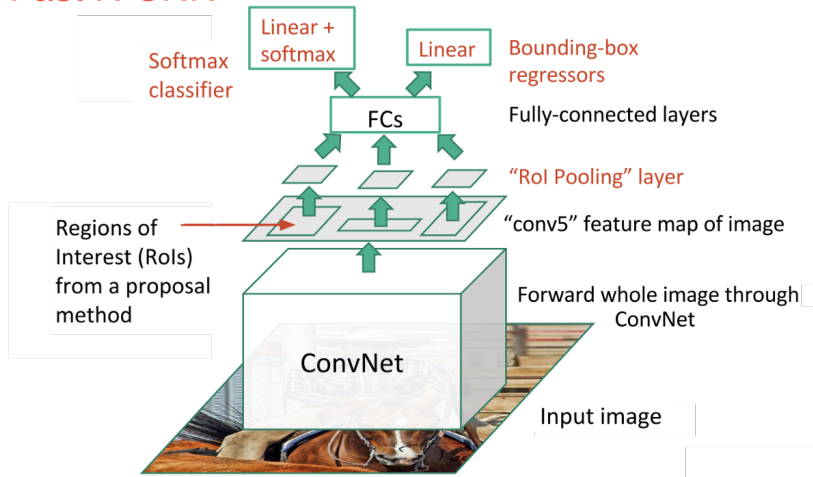
- 1 Escanea la imagen de entrada utilizando un algoritmo llamado *Selective Search*, y genera al rededor de 2000 regiones para probar si en éstas se encuentra algun objeto.
- 2 Ejecuta una red convolucional (capas convolucionales y de *max pooling* por cada una de las regiones.
- 3 Toma la salida de cada una de dichas CNNs.

Redes convolucionales basadas en regiones (RCNN)

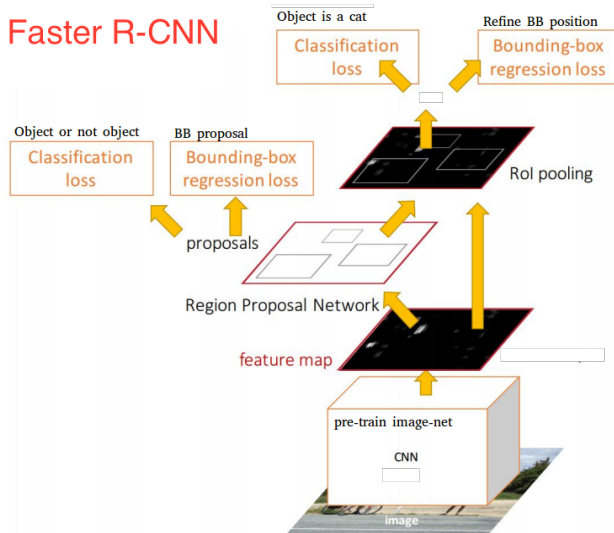


Redes convolucionales basadas en regiones mejorada (Fast-RCNN)

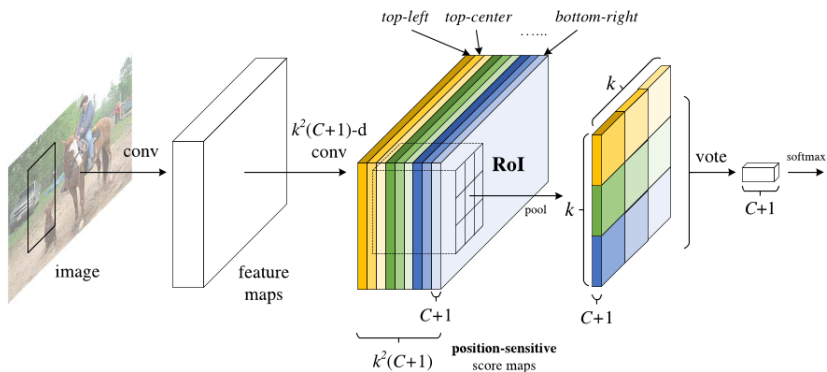
Fast R-CNN



Última version mejorada (Fast-RCNN)



Otra arquitectura no tan directa (RFCN)



Ejemplo de RFCN

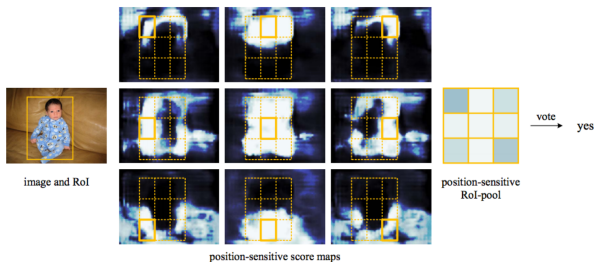
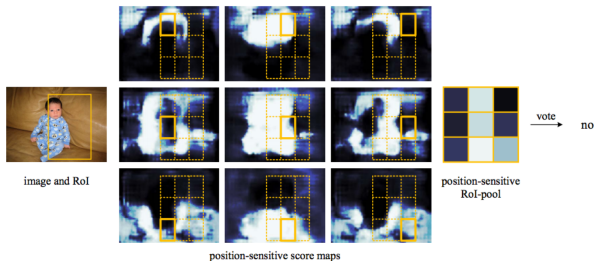
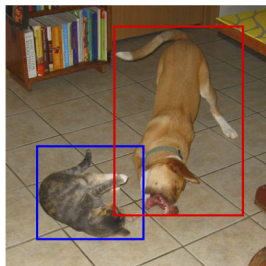


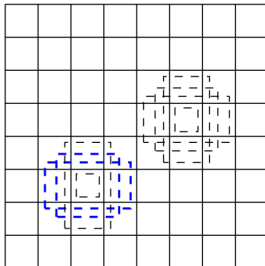
Figure 3: Visualization of R-FCN ($k \times k = 3 \times 3$) for the *person* category.



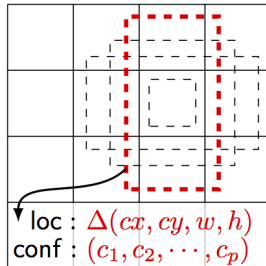
Ejemplo de SSD



(a) Image with GT boxes



(b) 8×8 feature map



(c) 4×4 feature map