

Fast R-CNN

Towards Real-Time Object Detection with Region Proposal Networks

A반 3조 배혜연

2021.05.30

- **Towards Real-Time Object Detection with Region Proposal Networks**
 - year : 2015
 - author : Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun
 - conference : NIPS

Problems of R-CNN

- 1) RoI (Region of Interest) 마다 CNN연산 -> 속도저하
- 2) multi-stage pipelines으로써 모델을 한번에 학습시키지 못함

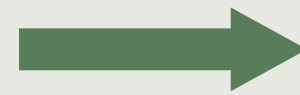


Fast R-CNN

Fast R-CNN

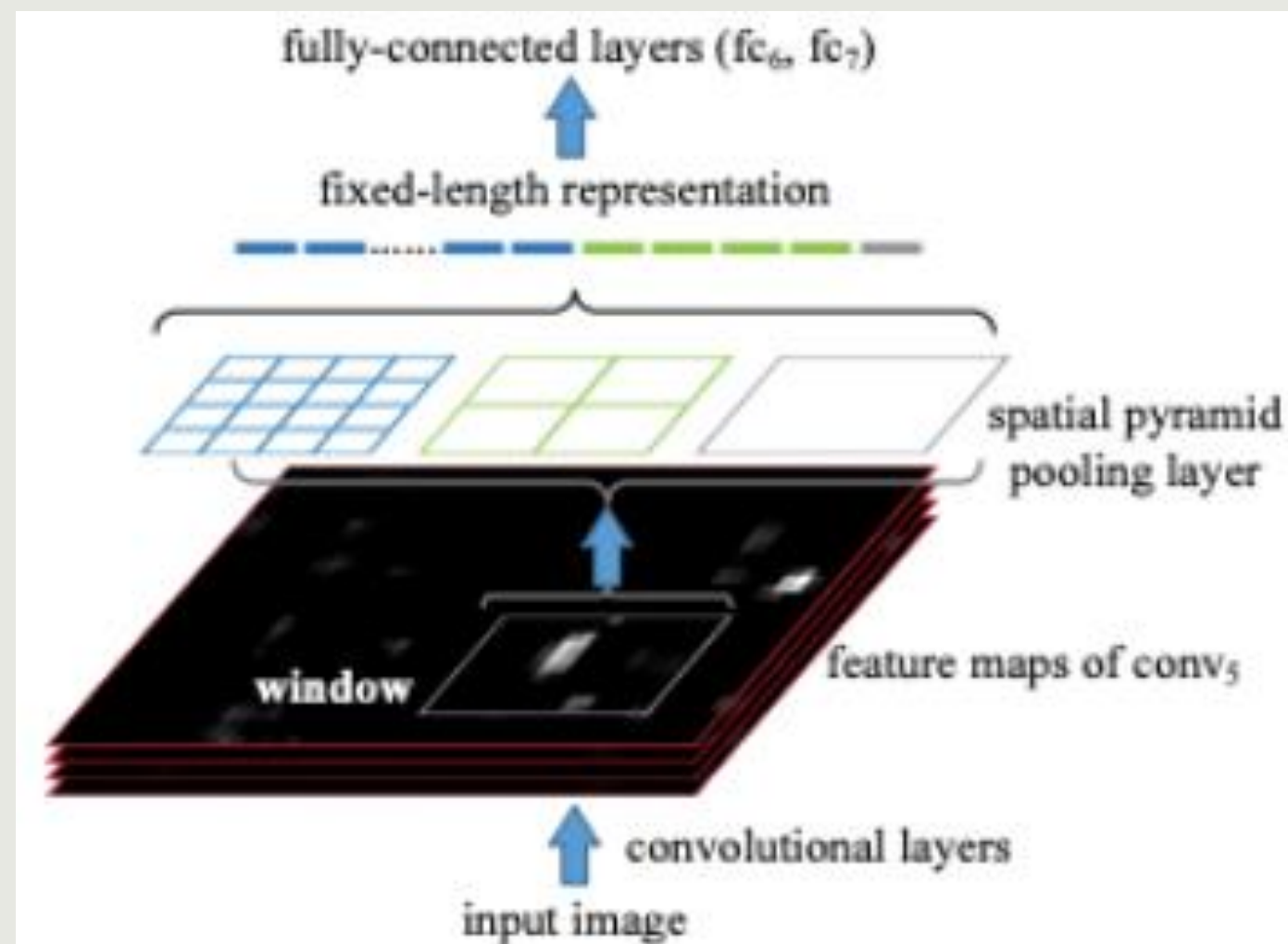
핵심 아이디어 : RoI Pooling

- CNN의 input : 입력 이미지 크기, 비율 관계없이 가능
- FC layer의 input: size를 맞춰줘야 함



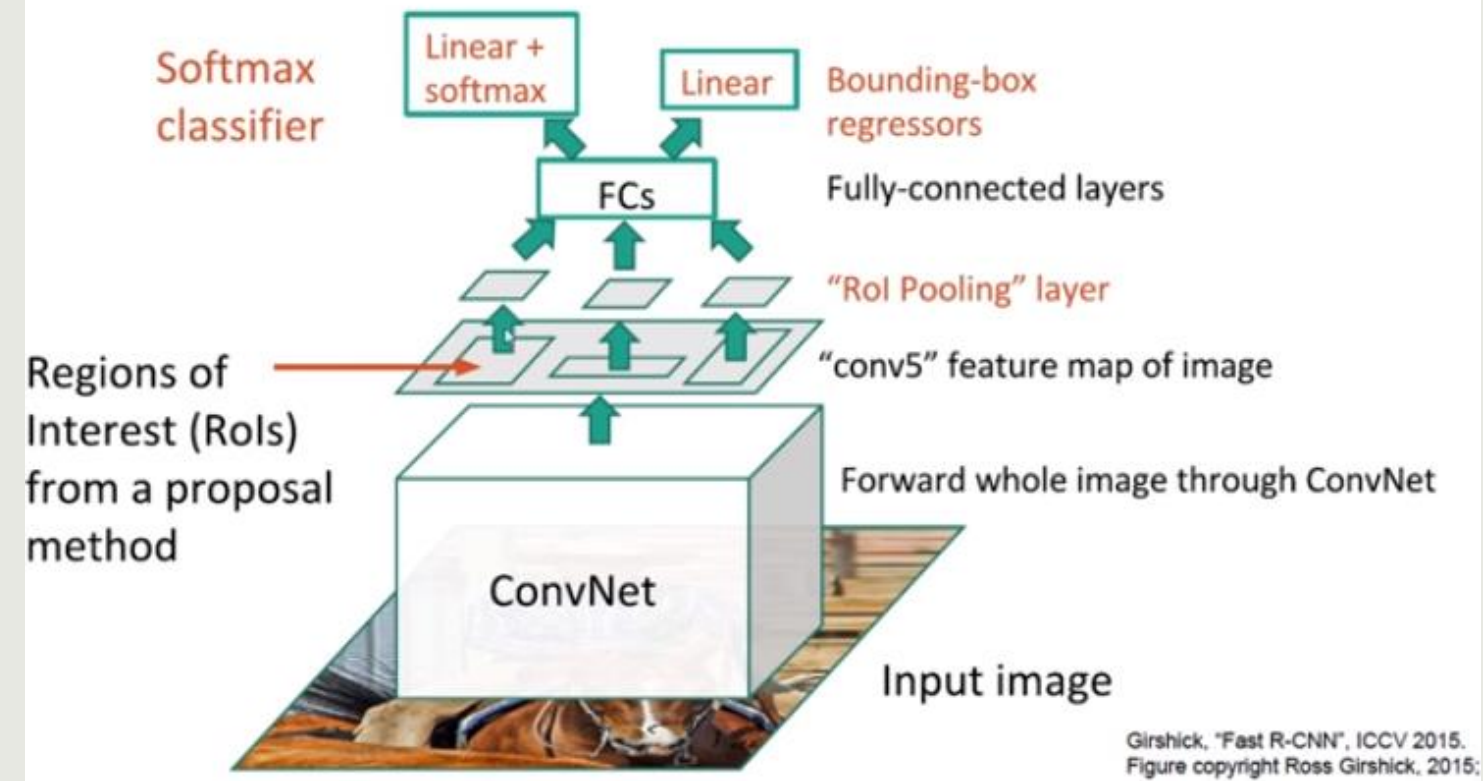
Spatial Pyramid Pooling(SPP) 제안

<Spatial Pyramid Pooling(SPP)>



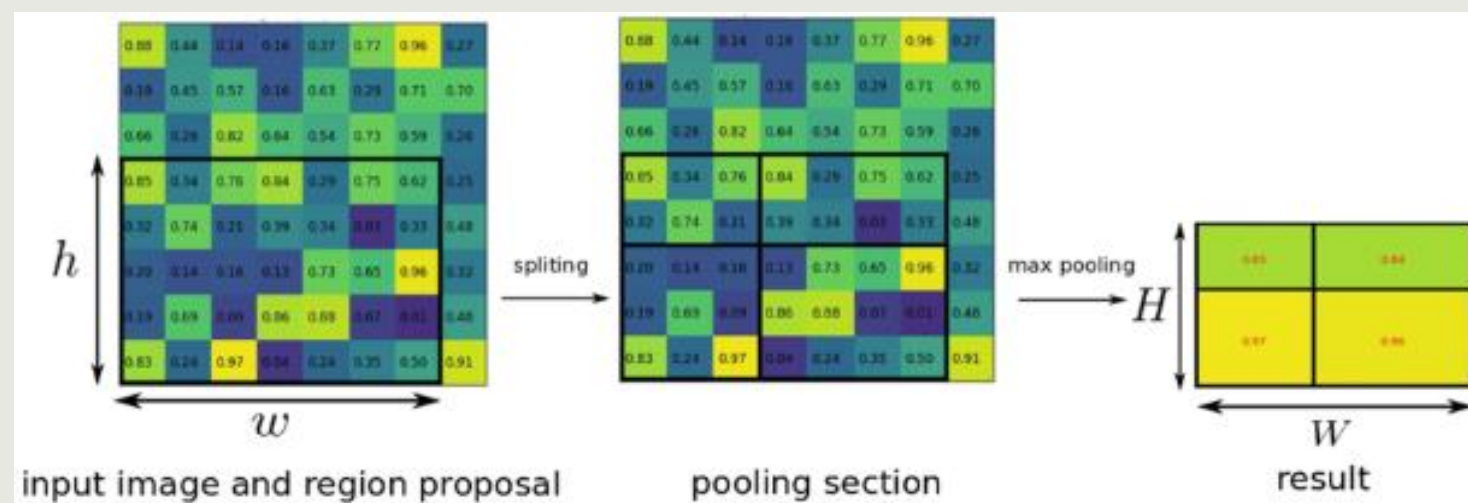
- 이미지를 CNN에 통과시켜 feature map을 추출
- 미리 정해진 4x4, 2x2, 1x1 영역의 피라미드로 feature map 분할
- bin내에서 max pooling을 적용하여 각 bin마다 하나의 값을 추출 (피라미드 한 칸을 bin이라 한다.)
- 각 피라미드 별로 뽑아낸 max값들을 이어붙여 고정된 크기 vector
-> FC layer의 input

Fast R-CNN



- 1개의 피라미드를 적용시킨 SPP로 구성
- Fast R-CNN에서 적용된 1개의 피라미드 SPP로 고정된 크기의 feature vector를 만드는 과정을 **"RoI Pooling"**

<RoI Pooling>

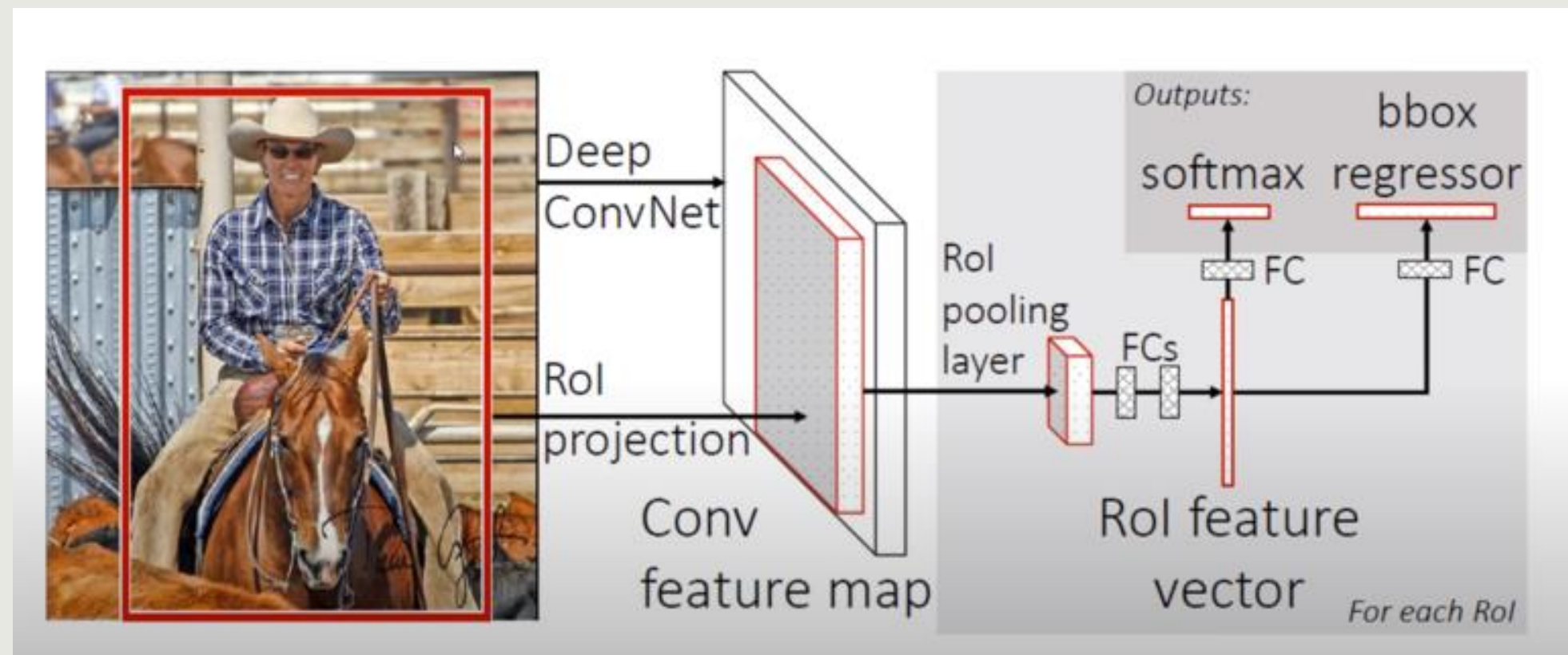


- 입력 이미지를 CNN에 통과시켜 feature map을 추출
- RoI(=region proposal)을 feature map에 projection
- RoI를 FC layer input 크기에 맞게 고정된 크기로 변형 가능

-> 속도 대폭 상승

Fast R-CNN

- 핵심은 RoI feature vector. Conv feature map의 마지막 부분에서 RoI pooling만 해서 RoI 부분만 추출하여 Fully Connected Layer로 넣는다.
그리고 최종적으로 물체에 대한 classification과 Bounding box regression 진행



- 동일 data가 각자 softmax(classification), bbox regressor(localization)으로 들어가기에 연산을 공유
- 모델이 end-to-end로 한 번에 학습시킬 수 있다



감사합니다