

Object Detection

POSTECH MIP Lab.

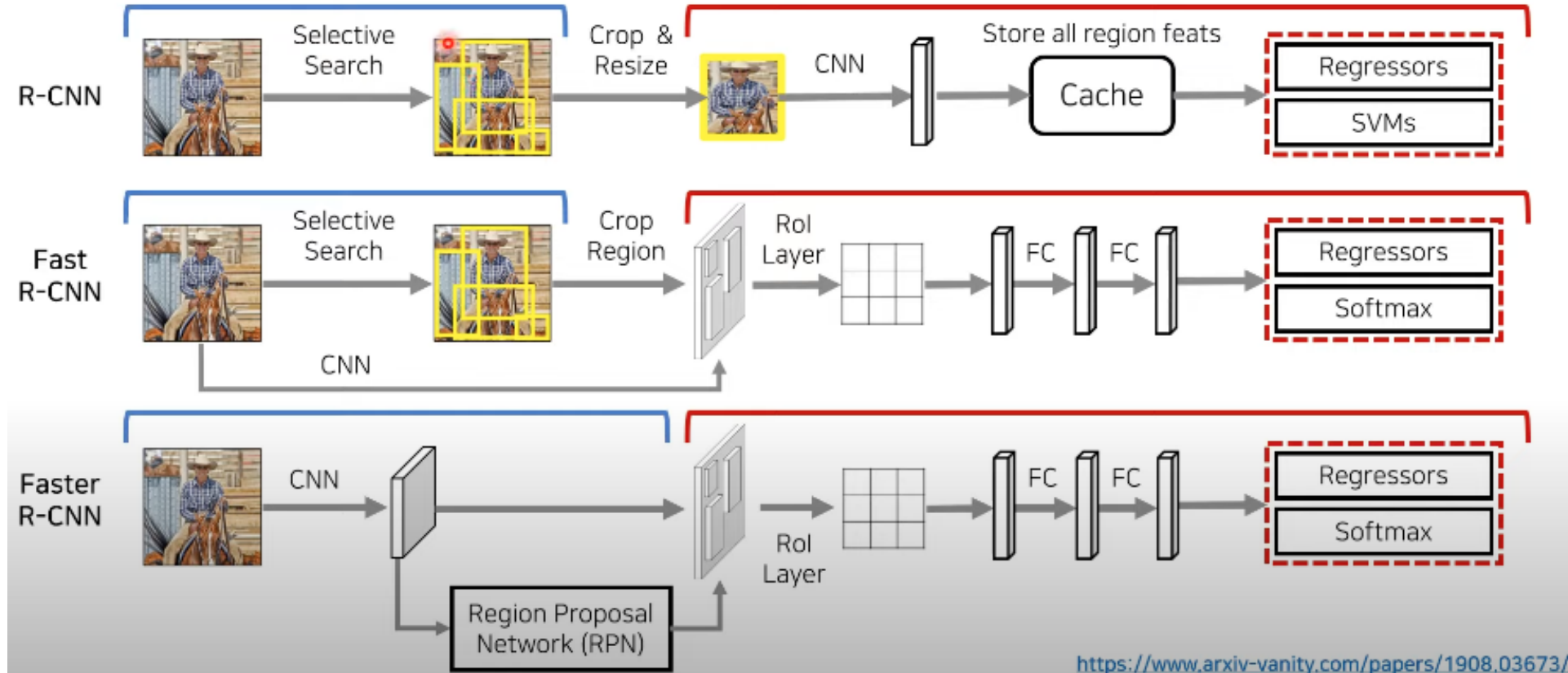
TA: Seunghun Baek, Joonhyuk Park, Soojin Hwang

Object Detection: A Naïve Approach

- Motivated by the great success of deep learning in image classification

$$\text{Object Detection} = \underbrace{\text{Box localization}}_{\text{Object proposals}} + \underbrace{\text{Box classification}}_{\text{CNN}}$$

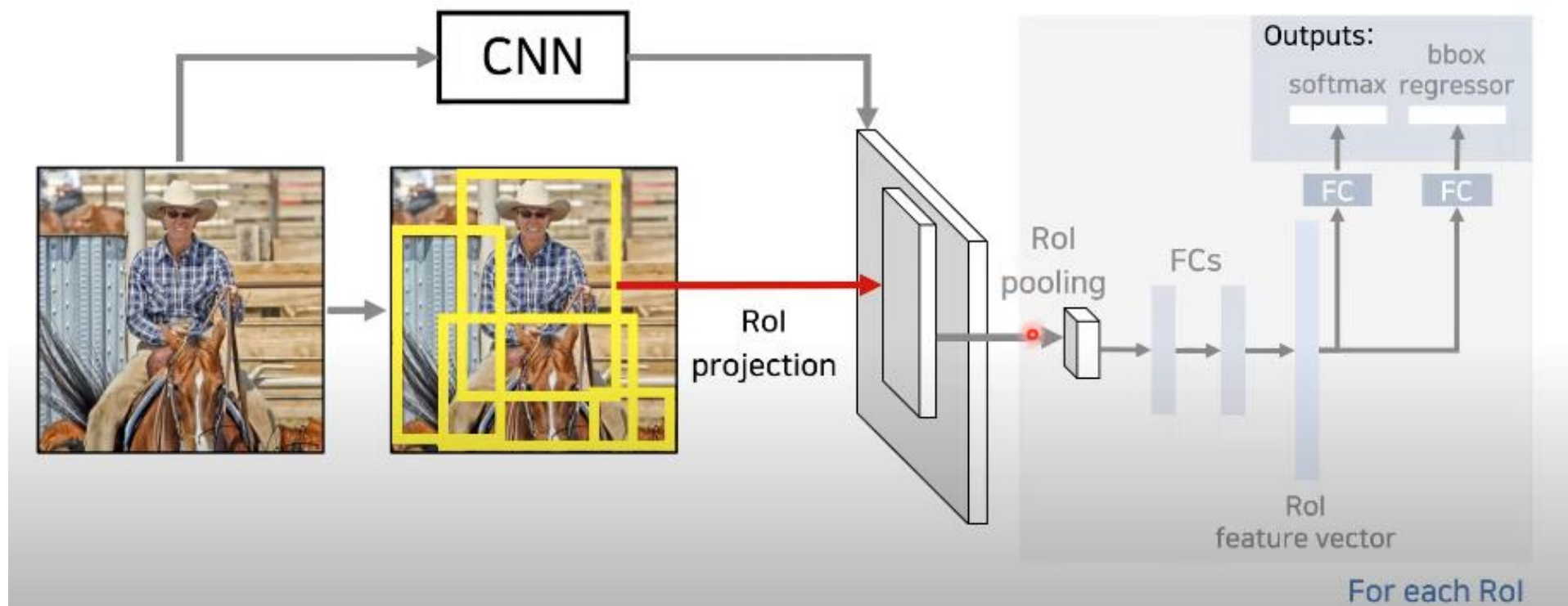
2-stage object detection



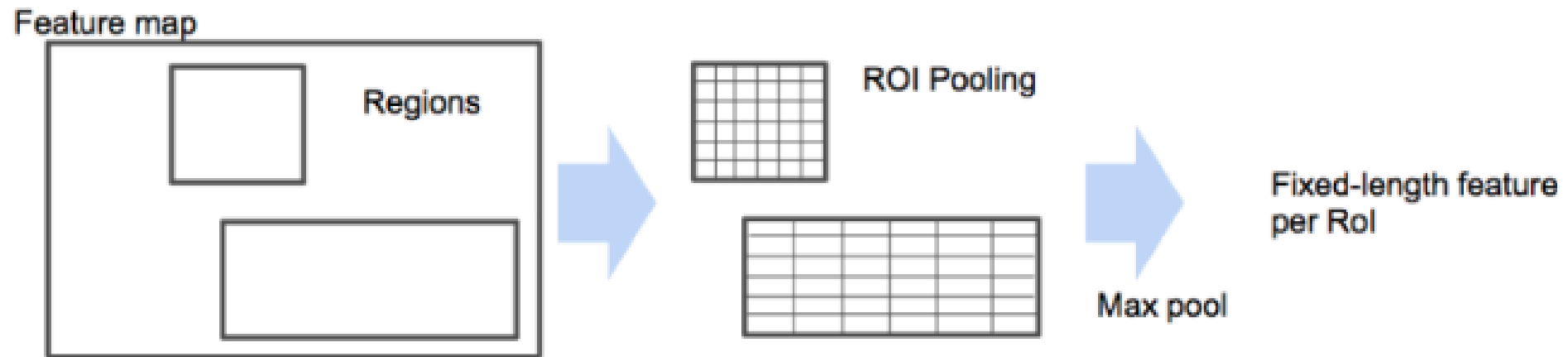
출처 : <https://www.youtube.com/watch?v=jqNCdjOB15s>

Fast R-CNN

동일한 Region proposal을 이용하되 이미지를 **한 번만 CNN**에 넣어 Feature Map 생성



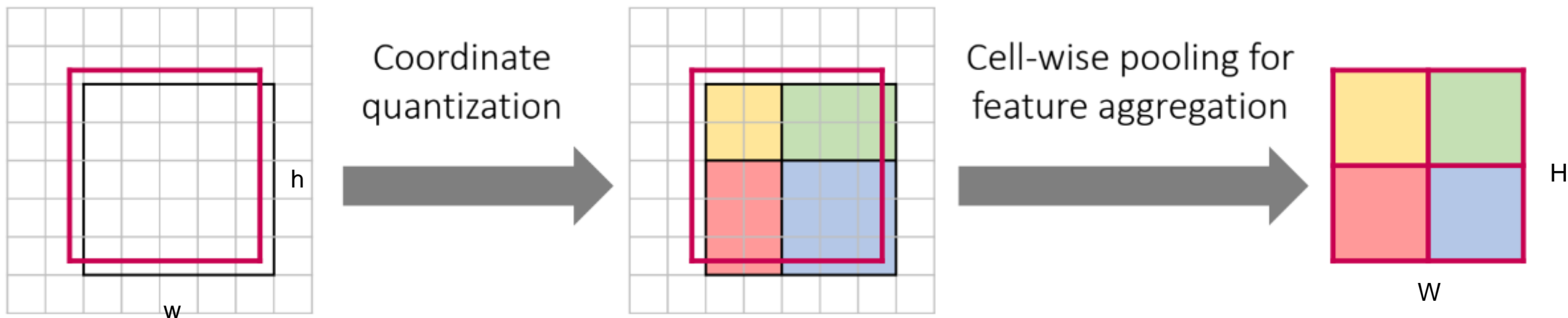
ROI pooling



Roi pooling layer의 pooling 과정

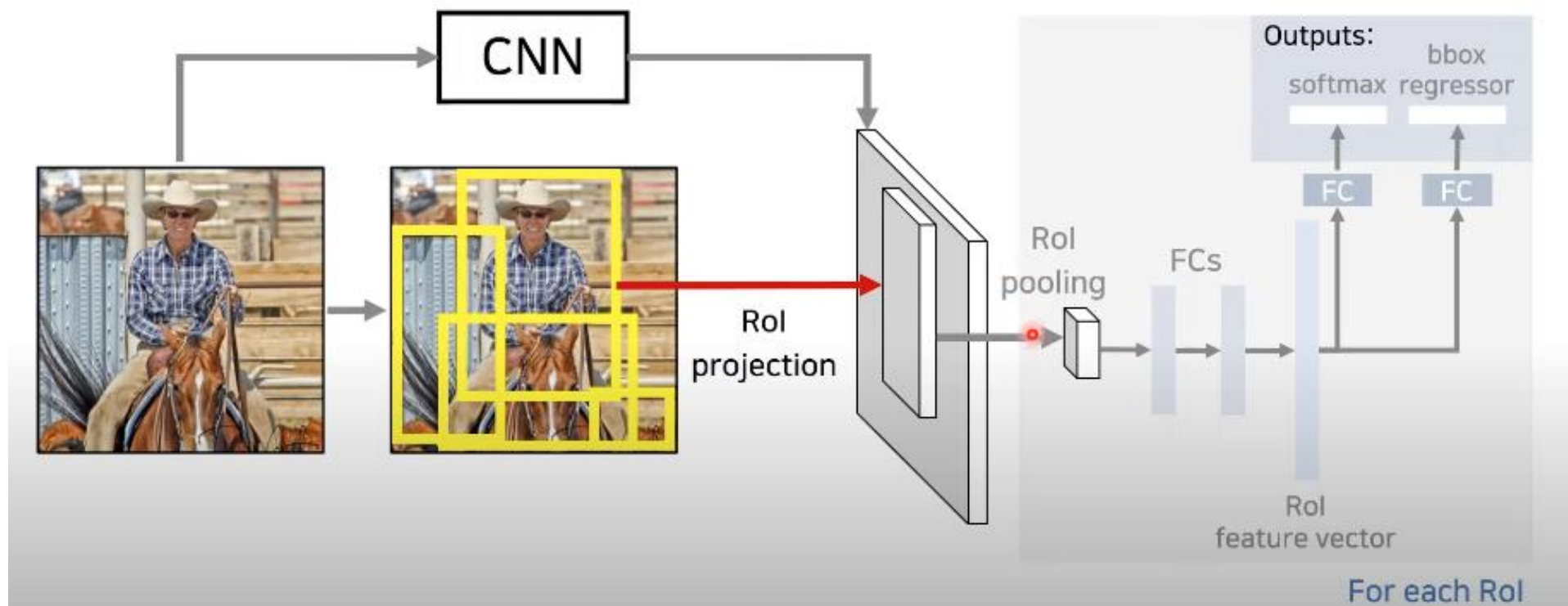
ROI pooling

- 각 ROI 영역에 대하여 max pooling을 이용해 고정된 크기의 벡터를 생성함.



Fast R-CNN

동일한 Region proposal을 이용하되 이미지를 **한 번만 CNN**에 넣어 Feature Map 생성



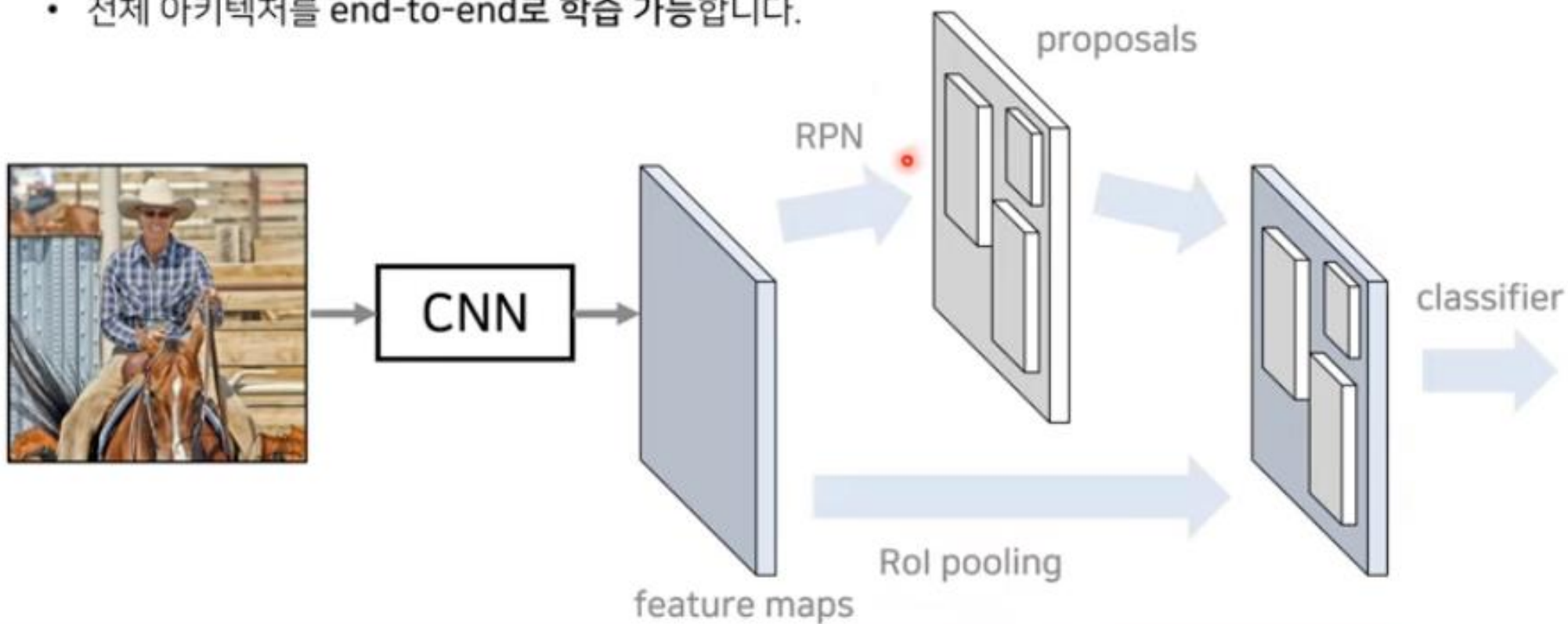
Fast R-CNN

- ~~Slow processing time~~ Much Faster By ROI-pooling, 1 forward path of CNN in one image
 - It needs to iterate forward propagation of input image patch over all proposals(~2000 forward propagations in practice)
- ~~Separate optimization of model components~~ End-to-End Learning (joint optimize all the component parameters)
 - Feature: CNN
 - ~~Classifier: SVM~~
 - Region proposal: Selective Search Window
 - Post-processing: Bounding box regression

→ It is not desirable to find optimal combination of all components

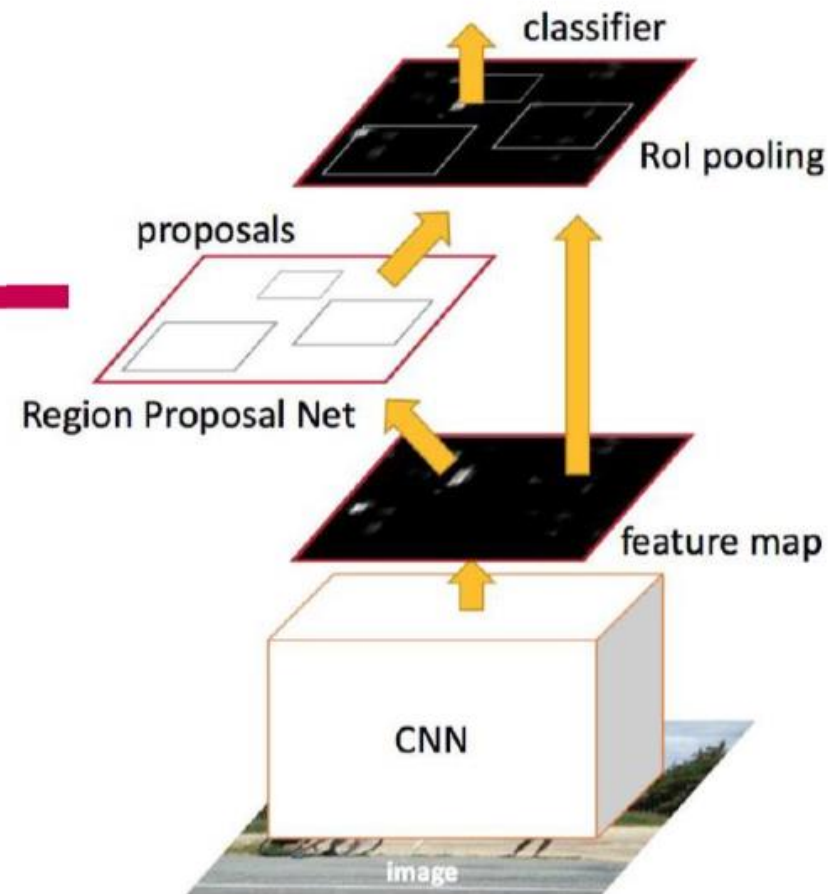
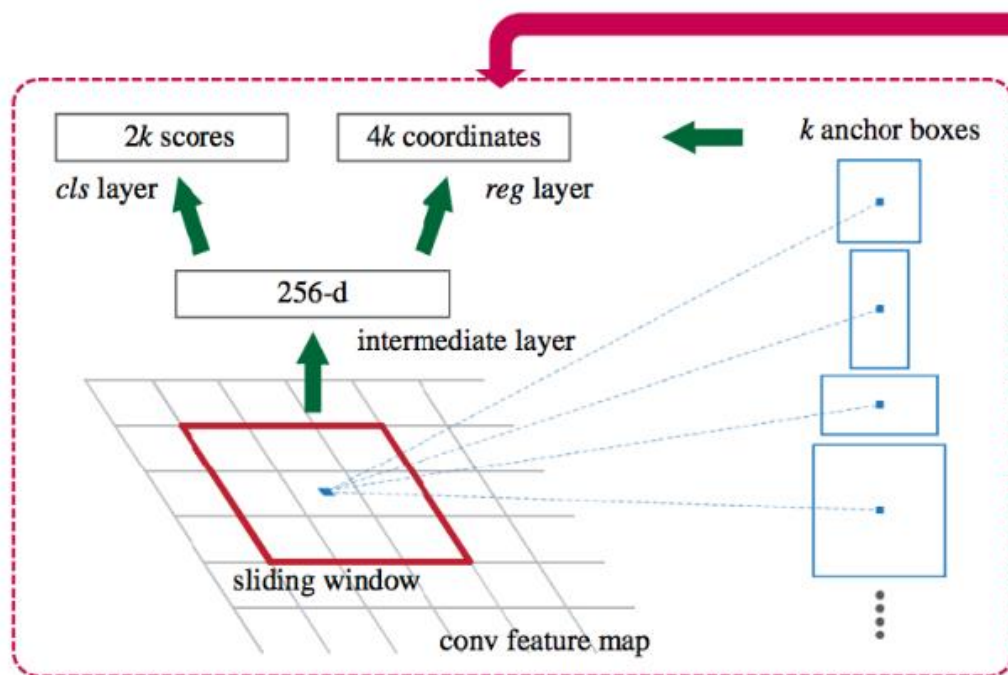
Faster R-CNN

- 병목(bottleneck)에 해당하던 Region Proposal 작업을 GPU 장치에서 수행하도록 합니다. (RPN 적용)
 - 전체 아키텍처를 end-to-end로 학습 가능합니다.



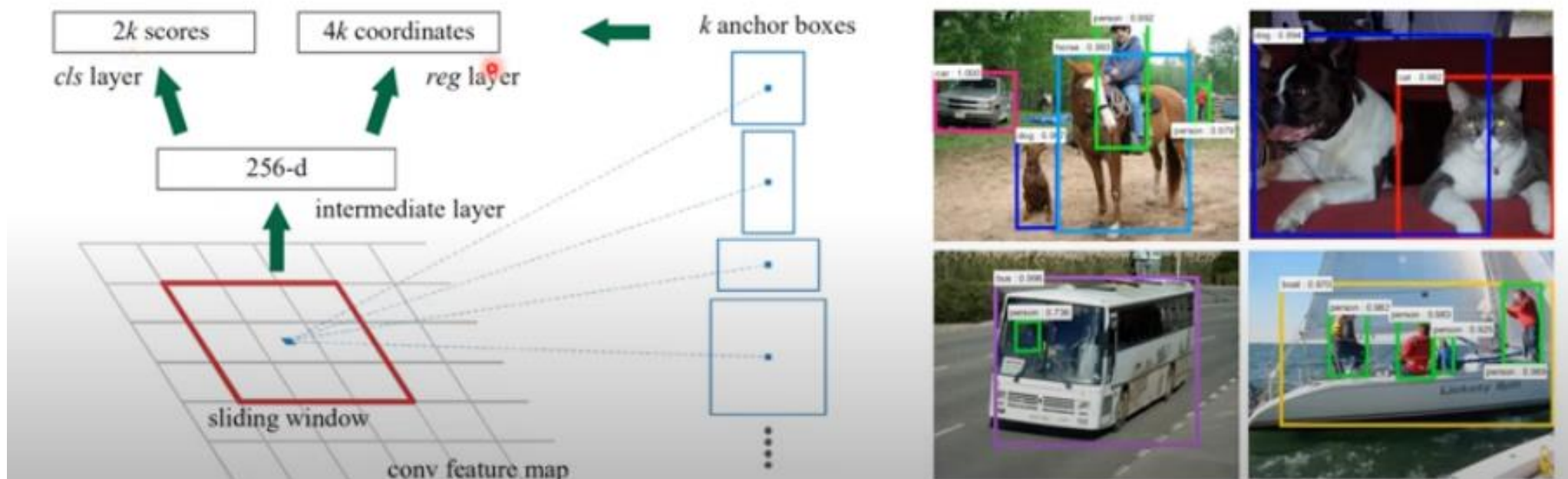
Faster R-CNN

- Fast R-CNN + Region Proposal Network (RPN)
 - Proposal computation into network
 - Multi-scaled reference boxes = Anchor

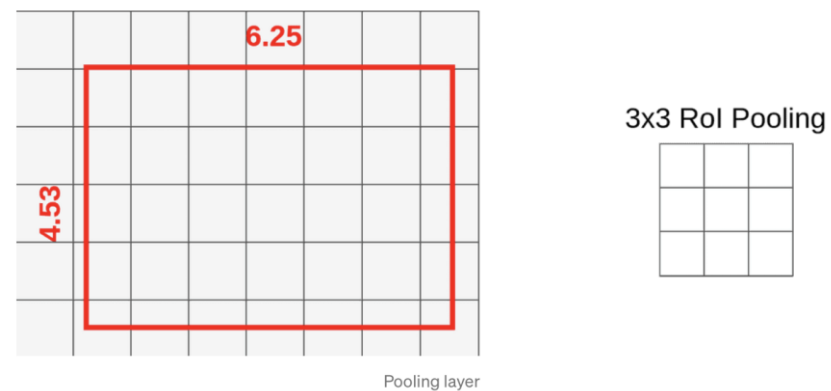
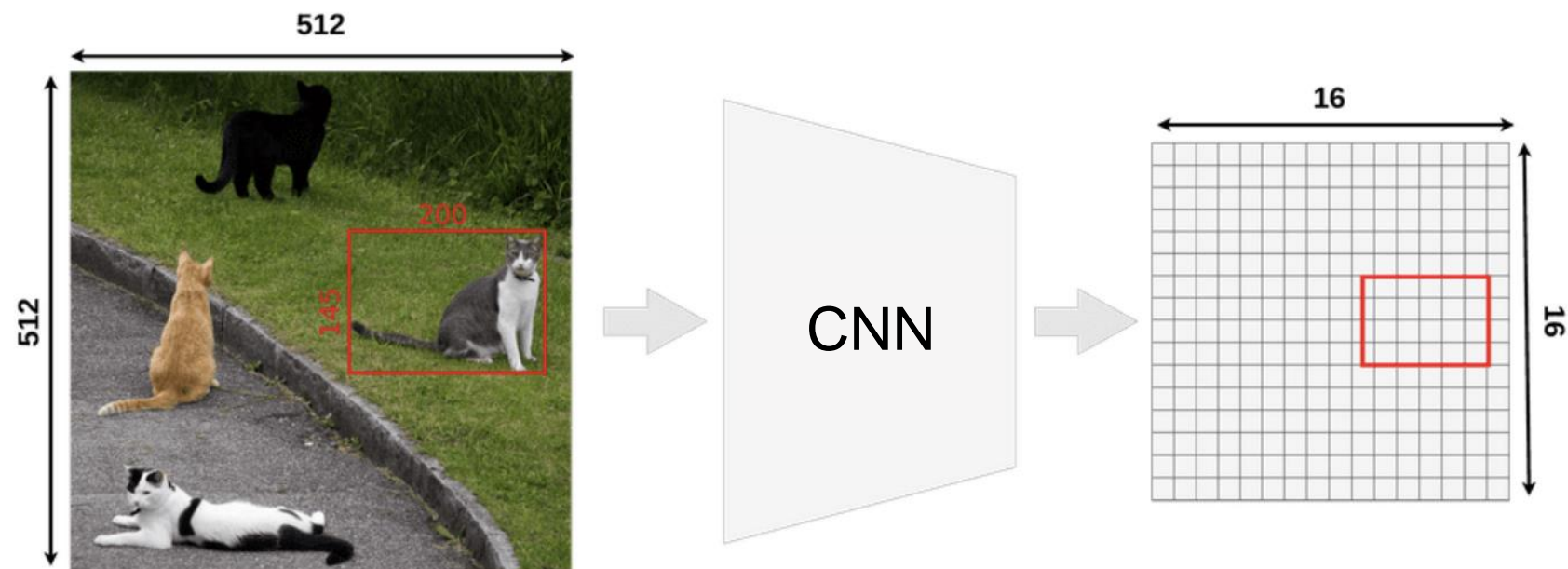


Region Proposal Networks (RPN)

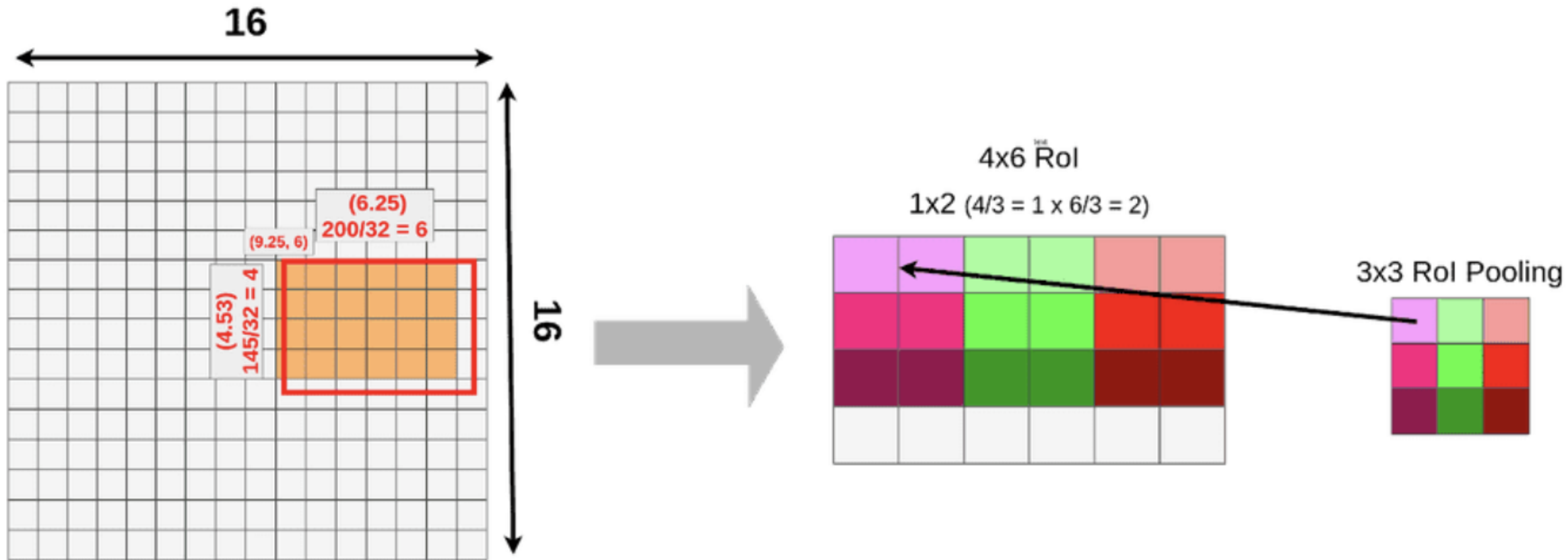
- RPN 네트워크는 feature map이 주어졌을 때 물체가 있을 법한 위치를 예측합니다.
 - k 개의 앵커 박스(anchor box)를 이용합니다.
 - 슬라이딩 윈도우(sliding window)를 거쳐 각 위치에 대해 Regression과 Classification을 수행합니다.



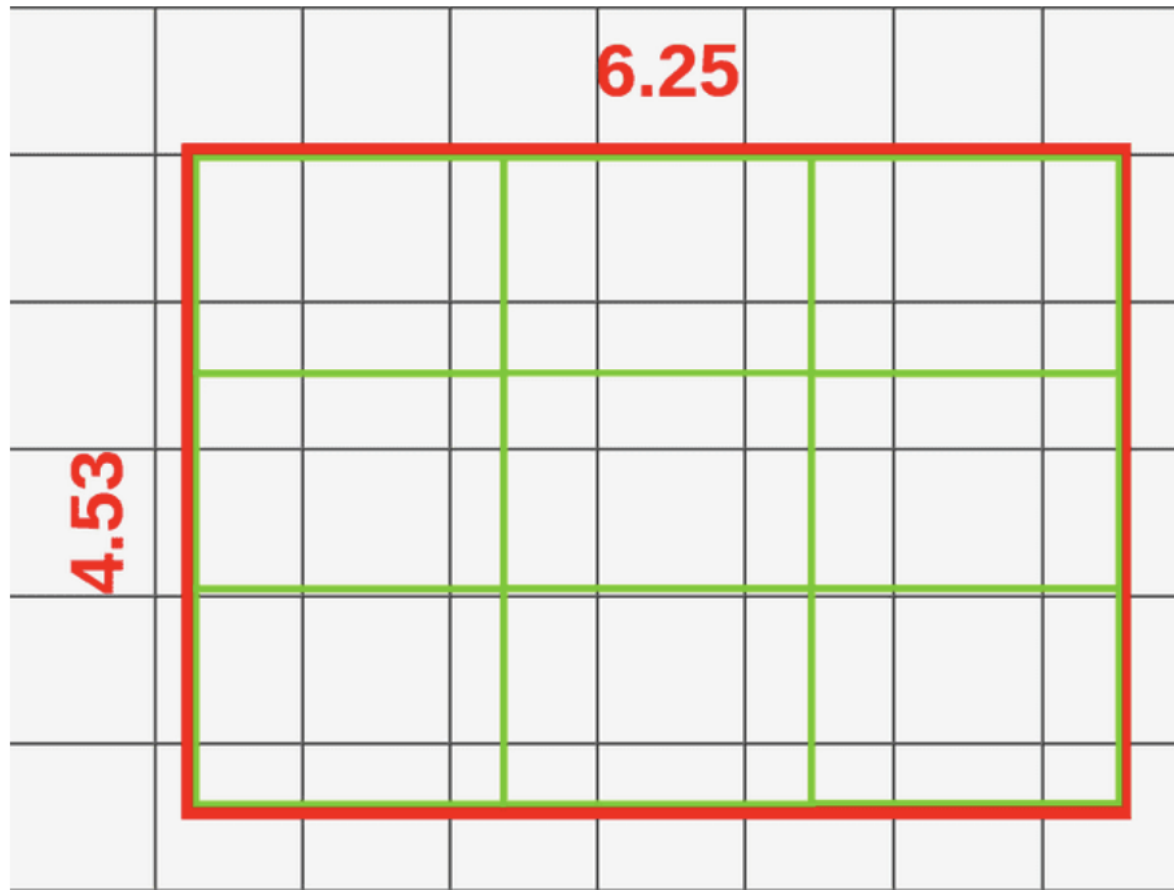
RoI Pooling



RoI Pooling

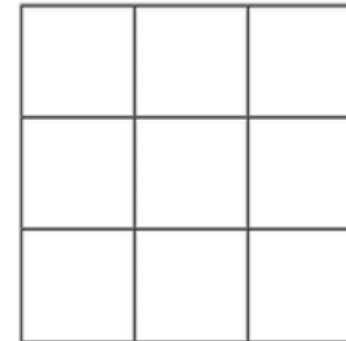


RoI Align

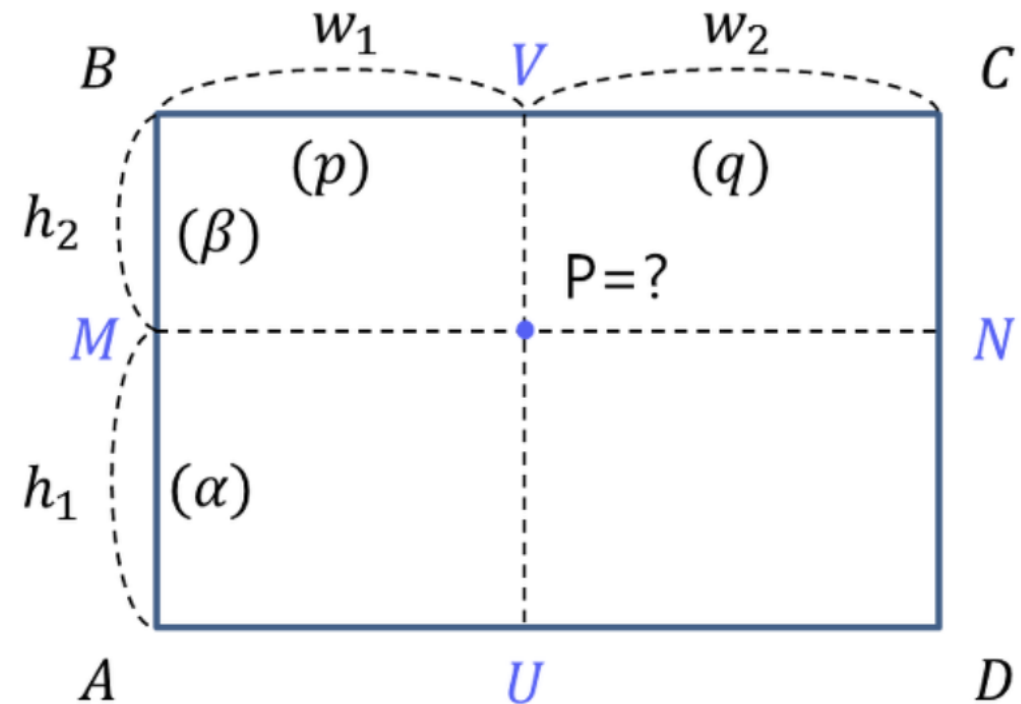


RoI divided into boxes

3x3 RoI Pooling



RoI Align



Get Features by Max Pooling

Faster R-CNN

- Result

per image

system	time	07 data	07+12 data
R-CNN	~50s	66.0	-
Fast R-CNN	~2s	66.9	70.0
Faster R-CNN	198ms	69.9	73.2

detection mAP on PASCAL VOC 2007, with VGG-16 pre-trained on ImageNet

R-CNN Variants들의 발전 방향 및 장단점 분석

발전 방향 ↓	R-CNN (CVPR 2014)	장점	CNN을 이용해 각 Region의 클래스를 분류할 수 있습니다.
		단점	전체 프레임워크를 End-to-End 방식으로 학습할 수 없습니다. 따라서 Global Optimal Solution을 찾기 어렵습니다.
	Fast R-CNN (ICCV 2015)	장점	Feature Extraction, RoI Pooling, Region Classification, Bounding Box Regression 단계(step)를 모두 End-to-End로 묶어서 학습할 수 있습니다.
		단점	여전히 첫 번째 Selective Search는 CPU에서 수행되므로 속도가 느립니다.
	Faster R-CNN (NIPS 2015)	장점	RPN을 제안하여, 전체 프레임워크를 End-to-End로 학습할 수 있습니다.
		단점	여전히 많은 컴포넌트로 구성되며, Region Classification 단계에서 각 특징 벡터(feature vector)는 개별적으로 FC layer로 Forward 됩니다.

Faster R-CNN실습