** Architecture of FASTER-RCNN with ResNet-50 Backbone and Region Proposal Network (RPN) **

Conv1 (7×7, 64 filters)
Conv2 (3 Bottlenecks)
Conv3 (4 Bottleneck)
Conv4 (6 Bottleneck)
Conv5 (3 Bottleneck)

Image RGB

Input Image

Convolution layers

(3×3)

Maxpooling

Sliding Window + 1 Conv. layer
Objectness Score

Maxpooling

(7×7)

ROI Align (feature map extraction) classification

Fully Connected Layers

2 layers

multi class-classification

Bounding Box Regressor

Region Proposal Network

RPN

CONV

For each spatial location
object classify

Bounding Box Regressor