

Input Image



YOLO

Faster
R-CNN

resize to 448x448

crops of 224x224

unified
conv.
layers

shared
conv.
layers

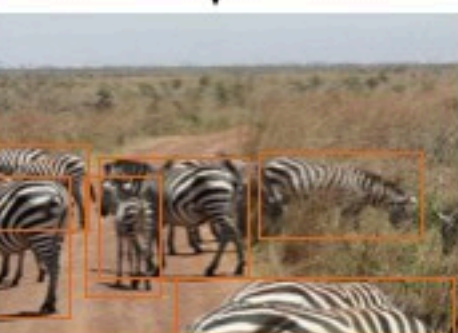
dense
layers

RPN
conv.
layers

2,000
RPN
proposals

raw output
98 boxes
classifications

dense
layers



Output Detections

raw output
300 boxes
classifications

Faster R-CNN is trained by
alternating optimization for
the RPN and dense layers