DETECTION
HEAD

INPUT

Image
416x416

CNN
BACKBONE

Extract
features

Box
x,y,w,h

Objectness
sigmoid

Classes
softmax

Dog: 95%
Box coords