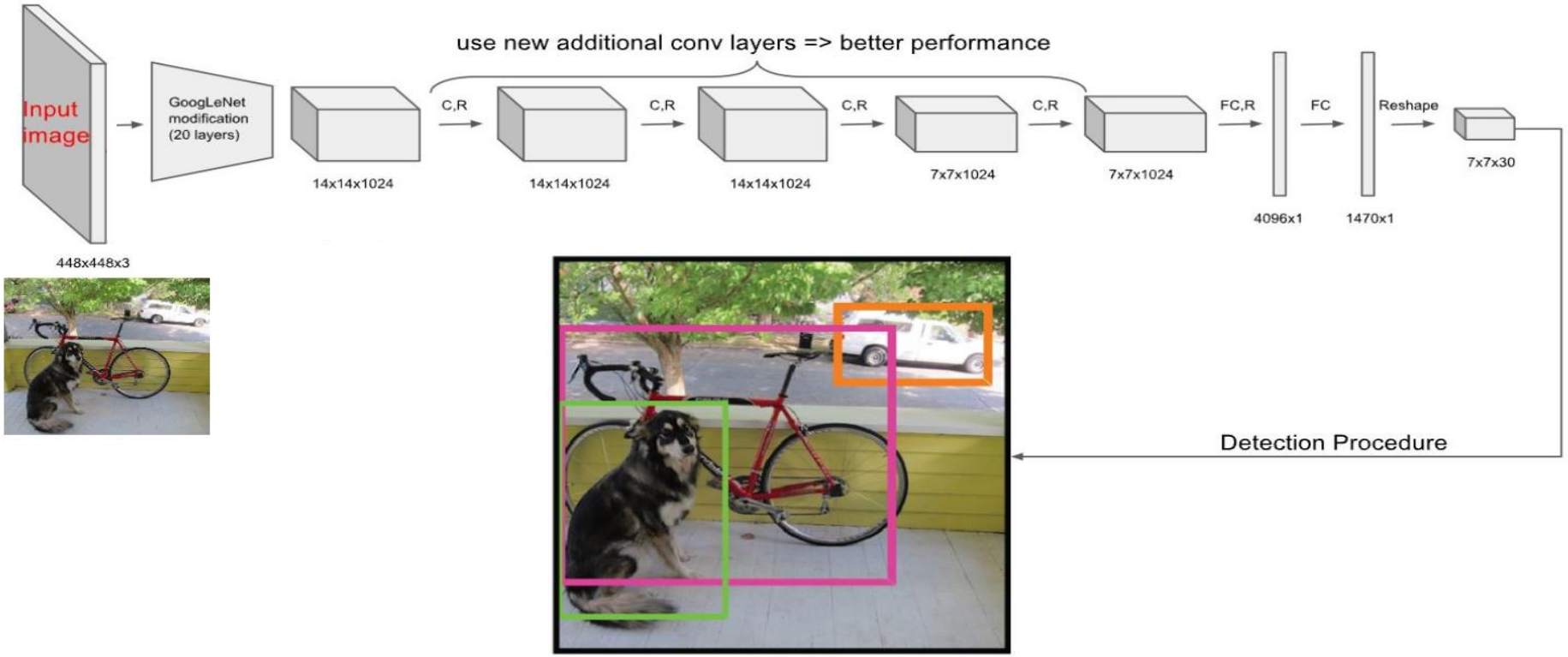
YOLO
Inference Process

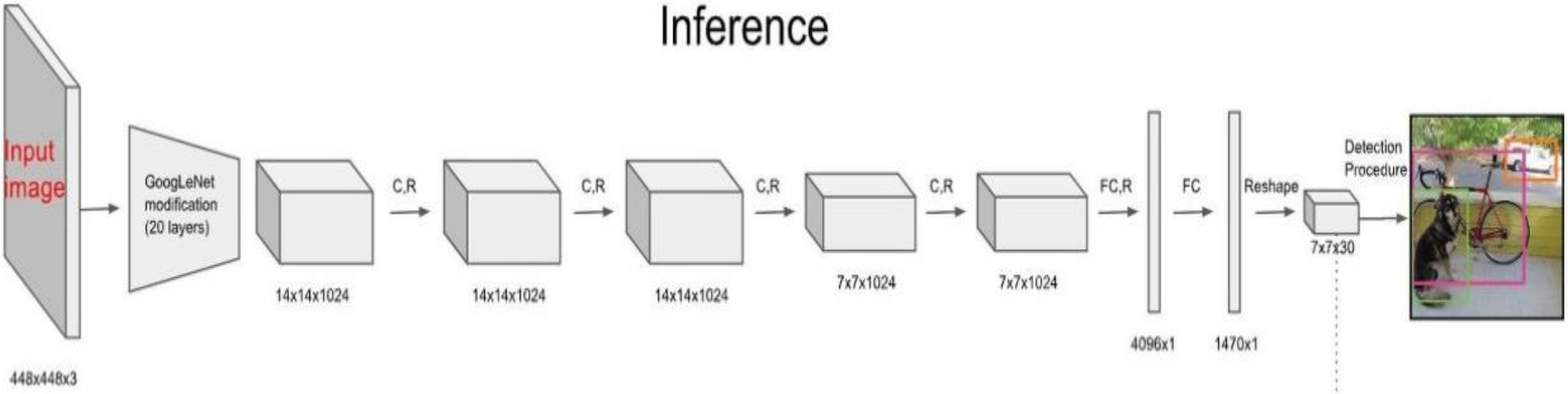
박성호 (neowizard2018@gmail.com)

YOLO Inference Process

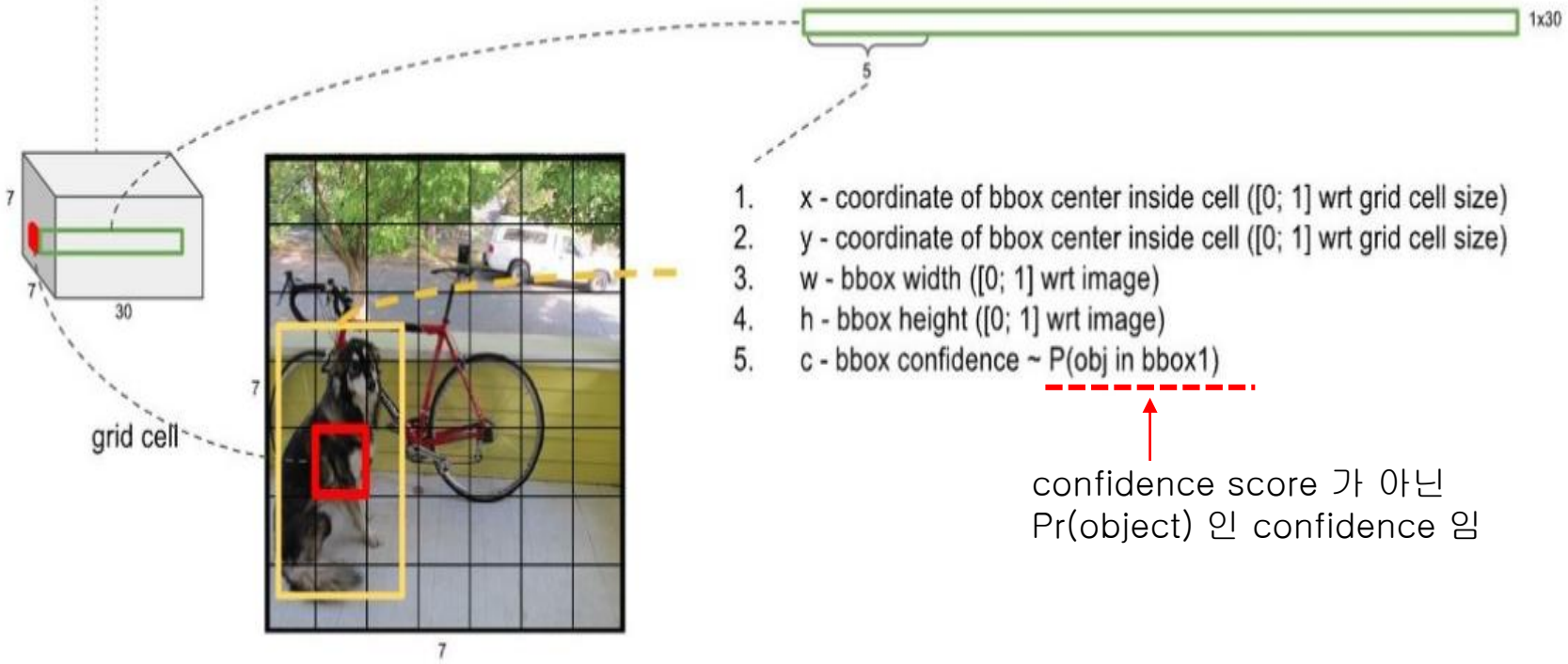
✓ 입력 이미지를 7 x 7 그리드 셀로 나누며, 각각의 그리드 셀에 들어있는 2개의 bounding box 정보와 물체의 클래스 정보가 들어있는 7x7x (5+5+20) 데이터가 YOLO 예측 결과임



1st bounding box of a grid cell



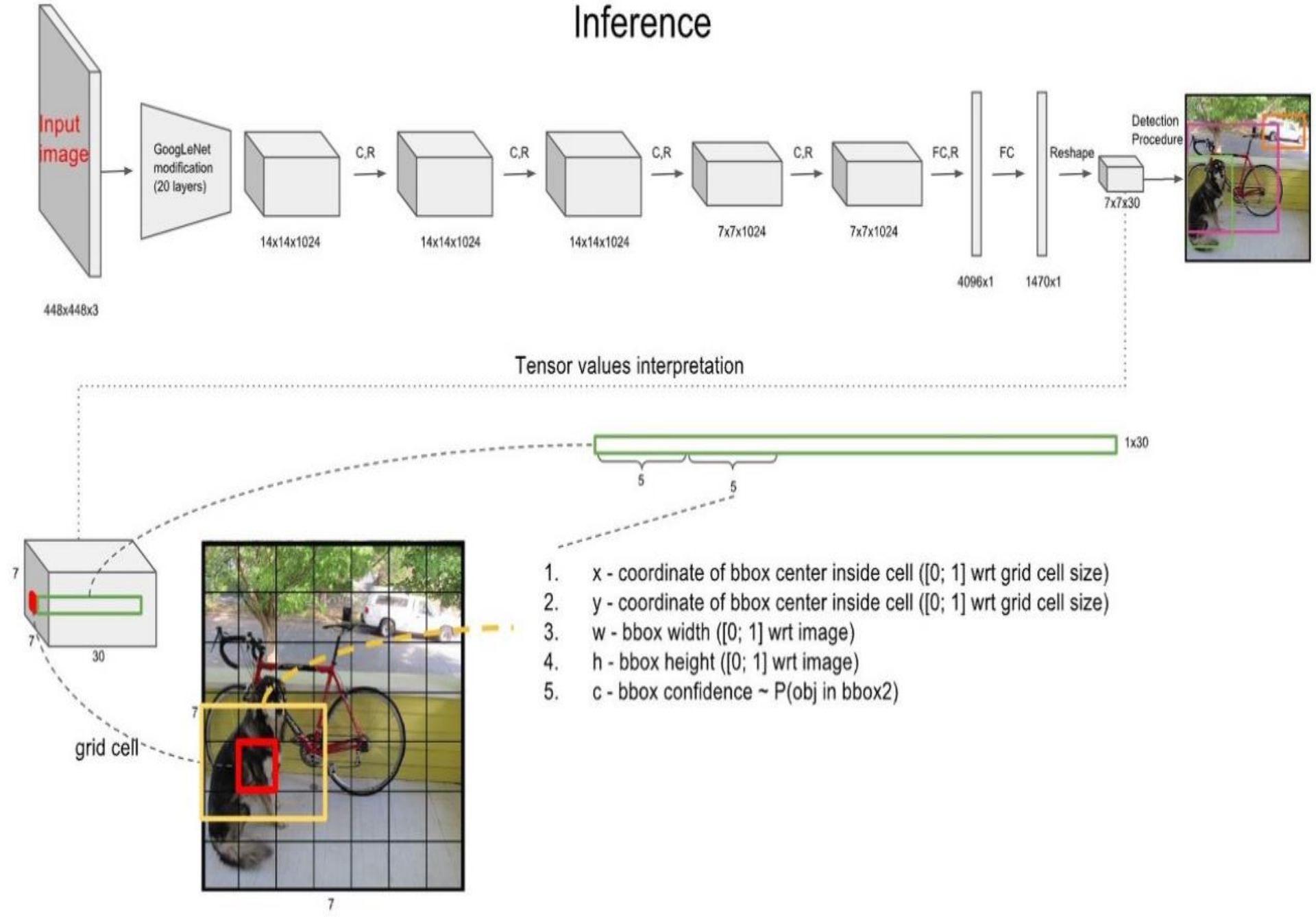
Tensor values interpretation



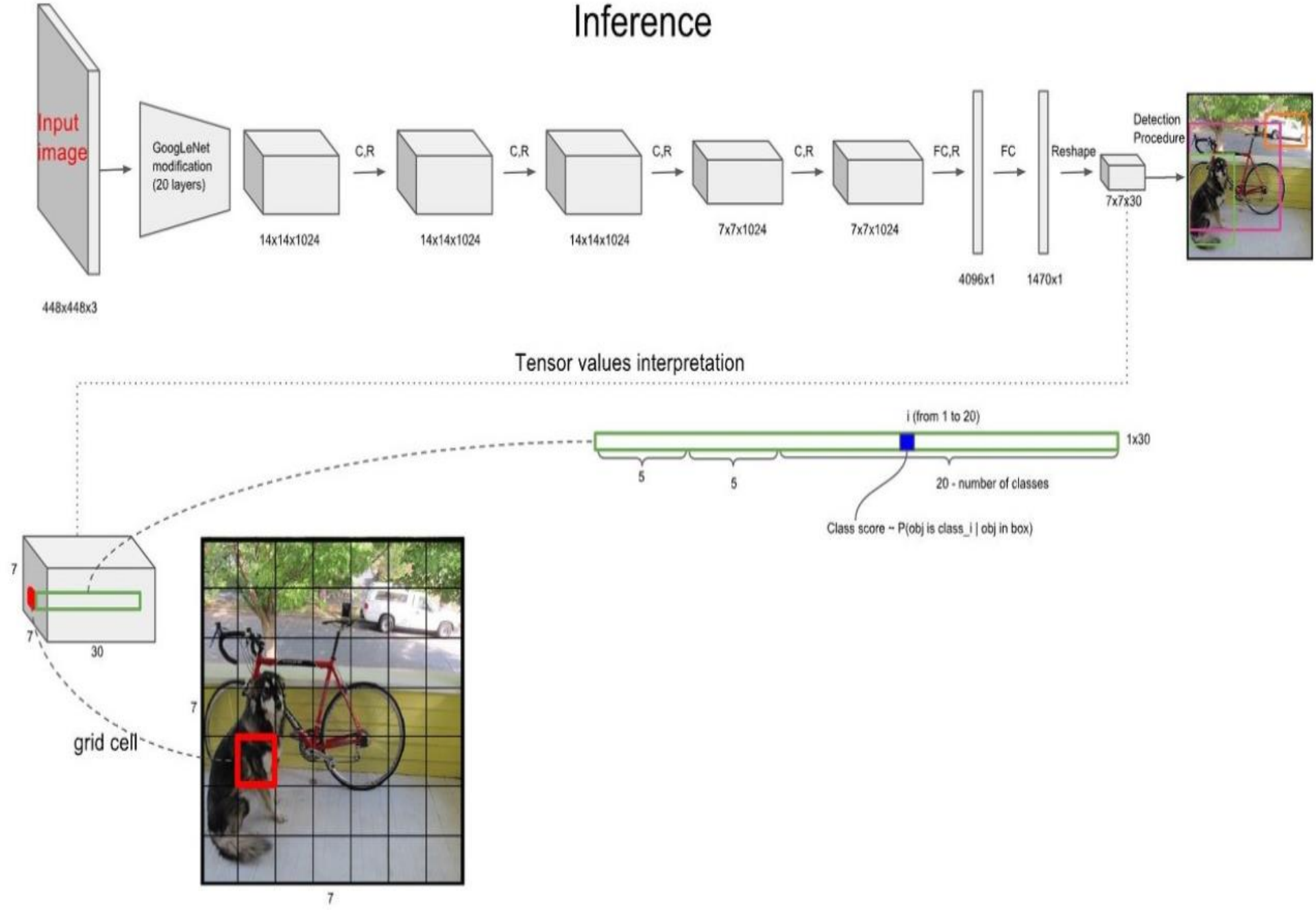
1. x - coordinate of bbox center inside cell ([0; 1] wrt grid cell size)
2. y - coordinate of bbox center inside cell ([0; 1] wrt grid cell size)
3. w - bbox width ([0; 1] wrt image)
4. h - bbox height ([0; 1] wrt image)
5. c - bbox confidence ~ P(obj in bbox1)

confidence score 가 아닌
Pr(object) 인 confidence 임

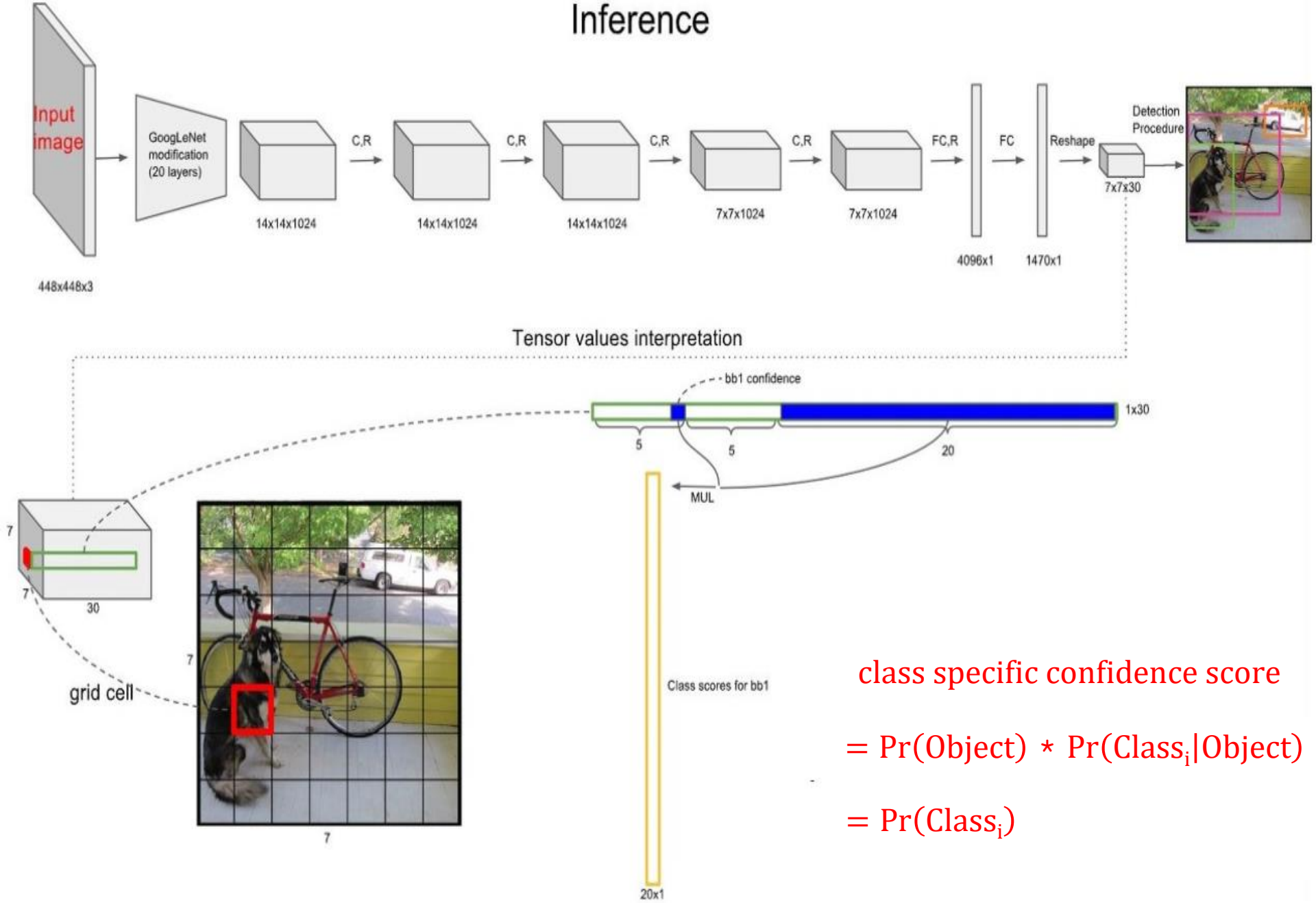
2nd bounding box of a grid cell



Class score


Inference
Input image
448x448x3
GoogLeNet modification (20 layers)
14x14x1024
C,R
14x14x1024
C,R
14x14x1024
C,R
7x7x1024
C,R
7x7x1024
FC,R
4096x1
FC
1470x1
Reshape
7x7x30
Detection Procedure
Tensor values interpretation
i (from 1 to 20)
1x30
5
5
20 - number of classes
Class score ~ P(obj is class_i | obj in box)
7
7
30
7
grid cell
7

각 grid cell 마다 2개의 class-specific confidence score 계산

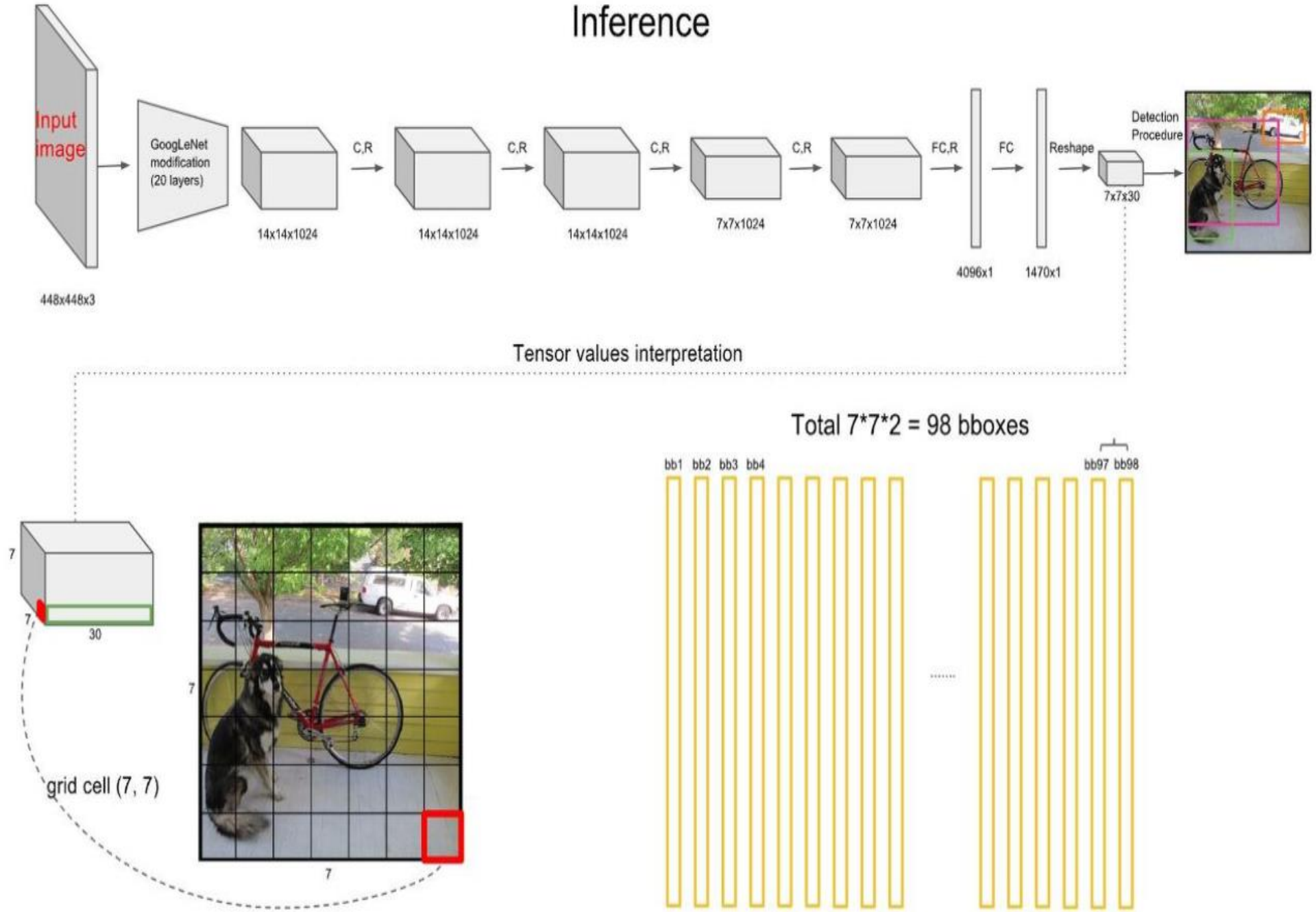


class specific confidence score

$= \Pr(\text{Object}) * \Pr(\text{Class}_i|\text{Object})$

$= \Pr(\text{Class}_i)$

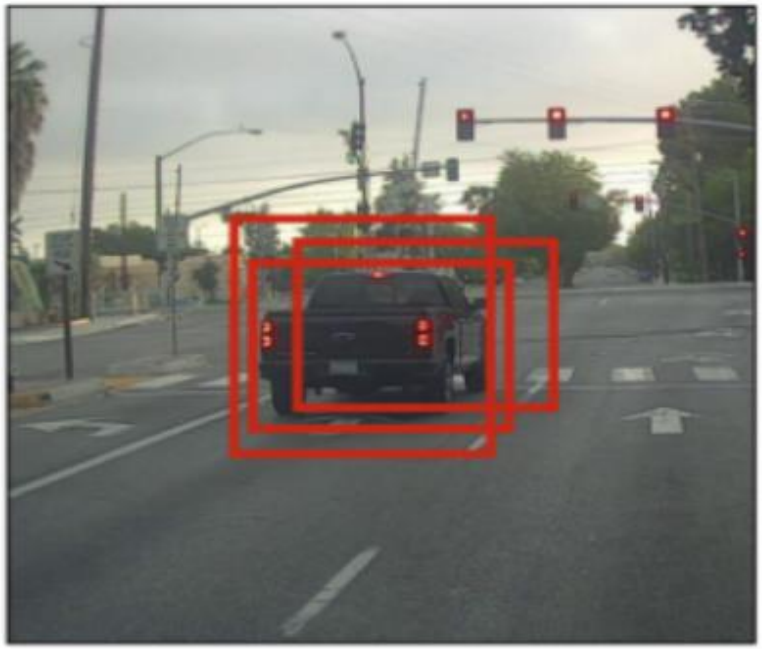
총 98개의 바운딩 박스의 class-specific confidence score 계산



➢ YOLO에서는 동일한 객체에 대하여 많은 bounding box가 잡힐 수 있음

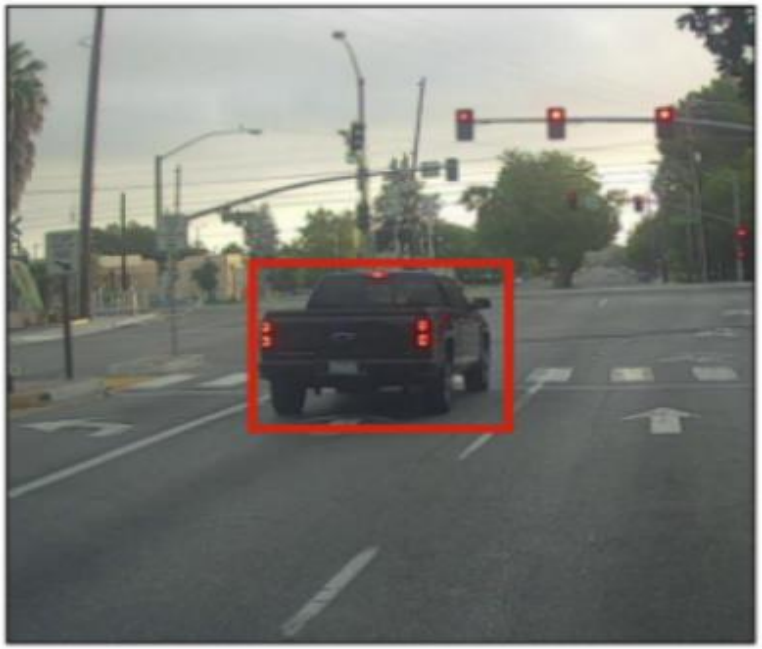
- 98개 bbox 각각이 가지고 있는 class specific confidence score에 대해서 각 20개의 클래스에 대해 신뢰도가 가장 높은 bbox만 남기고 나머지 bbox를 없애는 NMS (non-maximum suppression) 알고리즘을 이용하면, 객체에 대한 확률과 객체를 둘러싸고 있는 bbox 위치를 알아낼 수 있음.

before NMS



NMS

after NMS



※ YOLO 및 NMS 알고리즘 참고 사이트: https://goo.gl/byNZTn

deepsystems.io