

01

Real-Time Image Detection

1. YOLO
2. YOLOv4 - Architecture

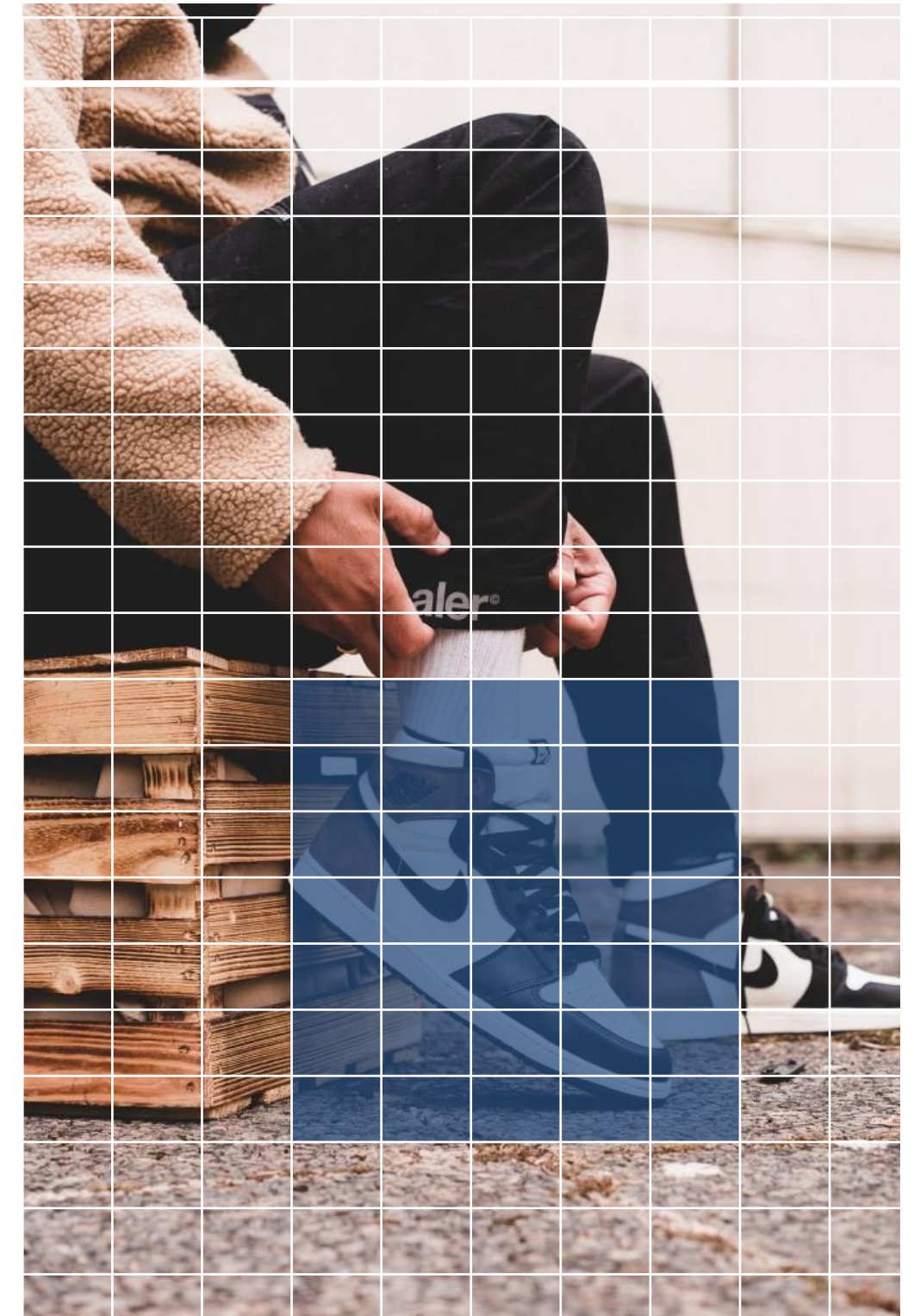
01

01. YOLO

02. YOLOv4

YOLO란? – You Only Look Once

1. 이미지 전체를 한번만 읽는다
 - 기존의 Image Detection Model들과의 차이점
 - 기존의 R-CNN, Fast-RCNN 등은 이미지를 분할 후 CNN으로 분석
2. 통합된 모델을 사용
 - 기존의 모델들은 다양한 전처리 함수와 인공신경망 모델을 결합
 - YOLO는 단순하고 통합된 모델을 사용해서 편리함
3. 기존의 모델들 보다 빨라진 실시간 이미지 인식
 - Faster R-CNN보다 약 6배 빠름



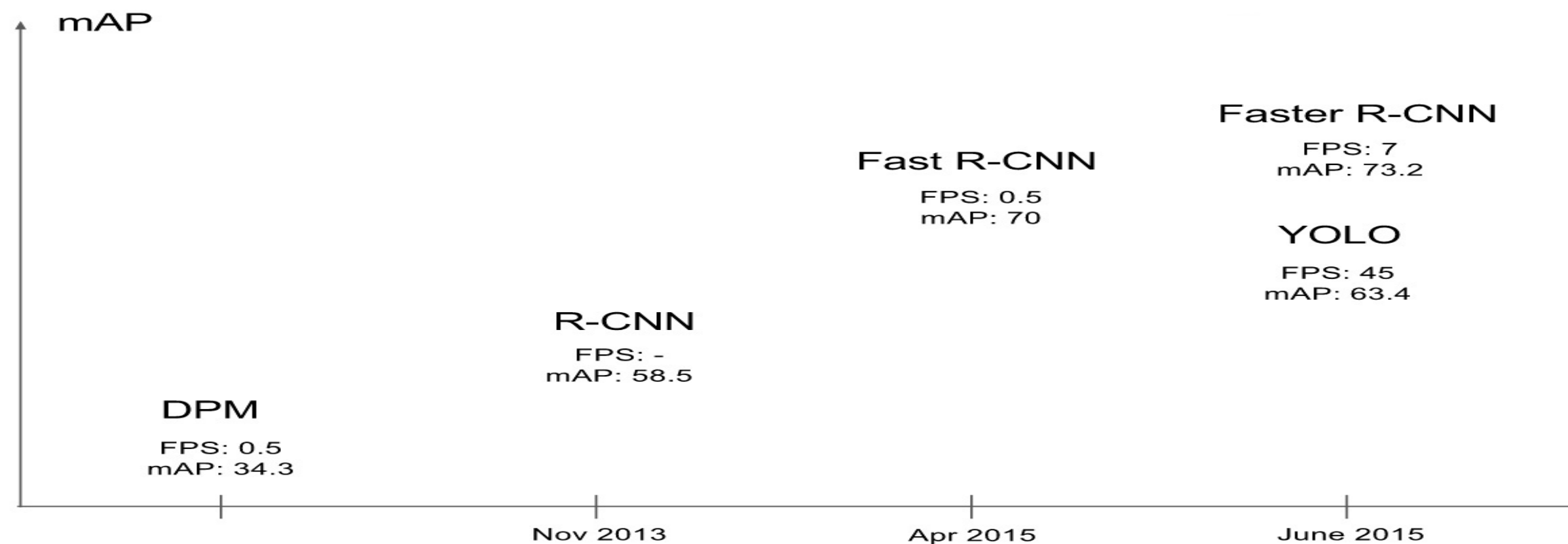
01

01. YOLO

02. YOLOv4

YOLO의 특징

- ▶ 간단한 처리 과정으로 속도가 매우 빠름
- ▶ 다른 모델에 비해 낮은 False-Positive를 보임
- ▶ 자연 이미지로 학습하여 다른 모델보다 높은 성능을 보임
- ▶ 다른 모델에 비해 상대적으로 낮은 정확도



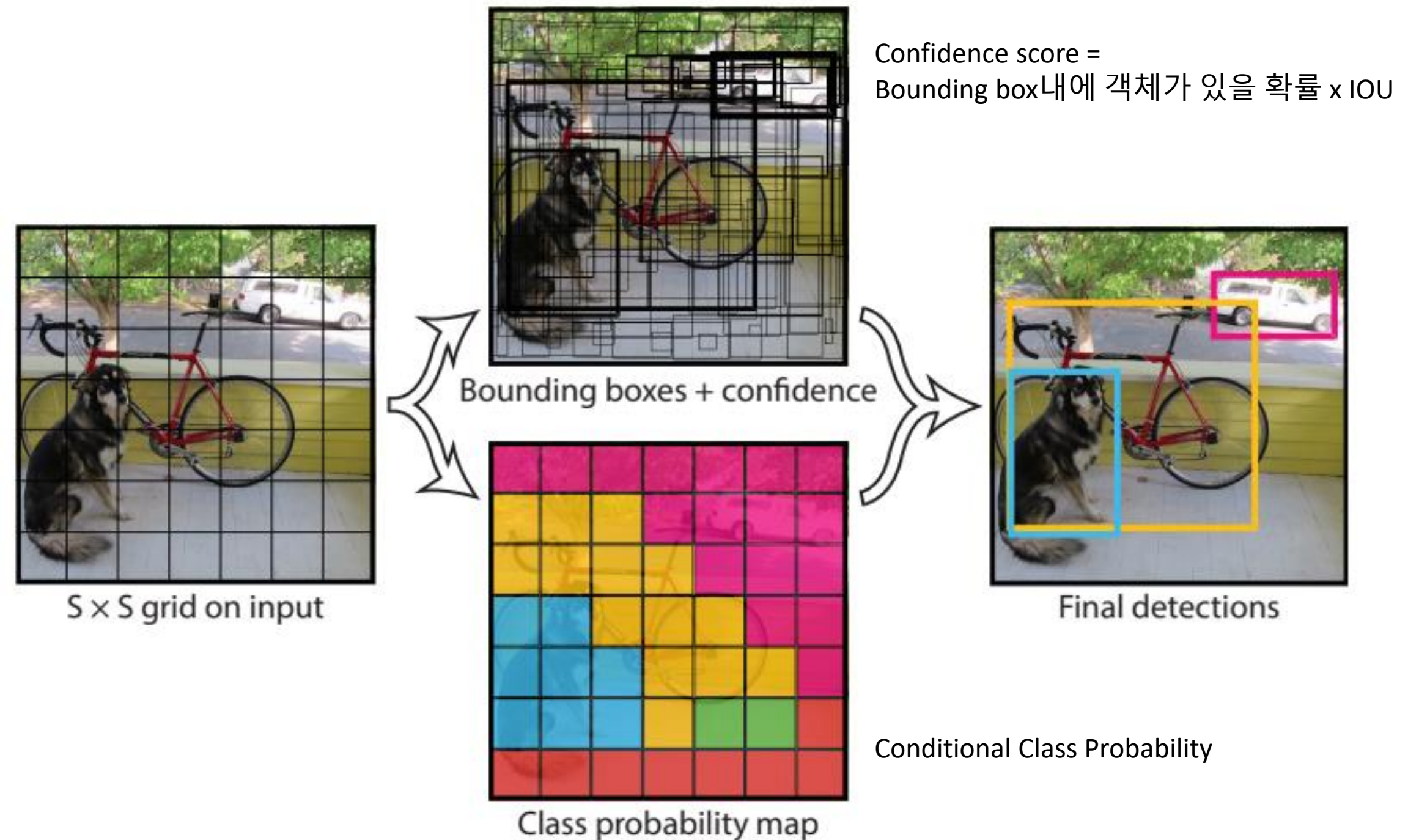
01

01. YOLO

02. YOLOv4

YOLO의 구조

Unified Detection



01

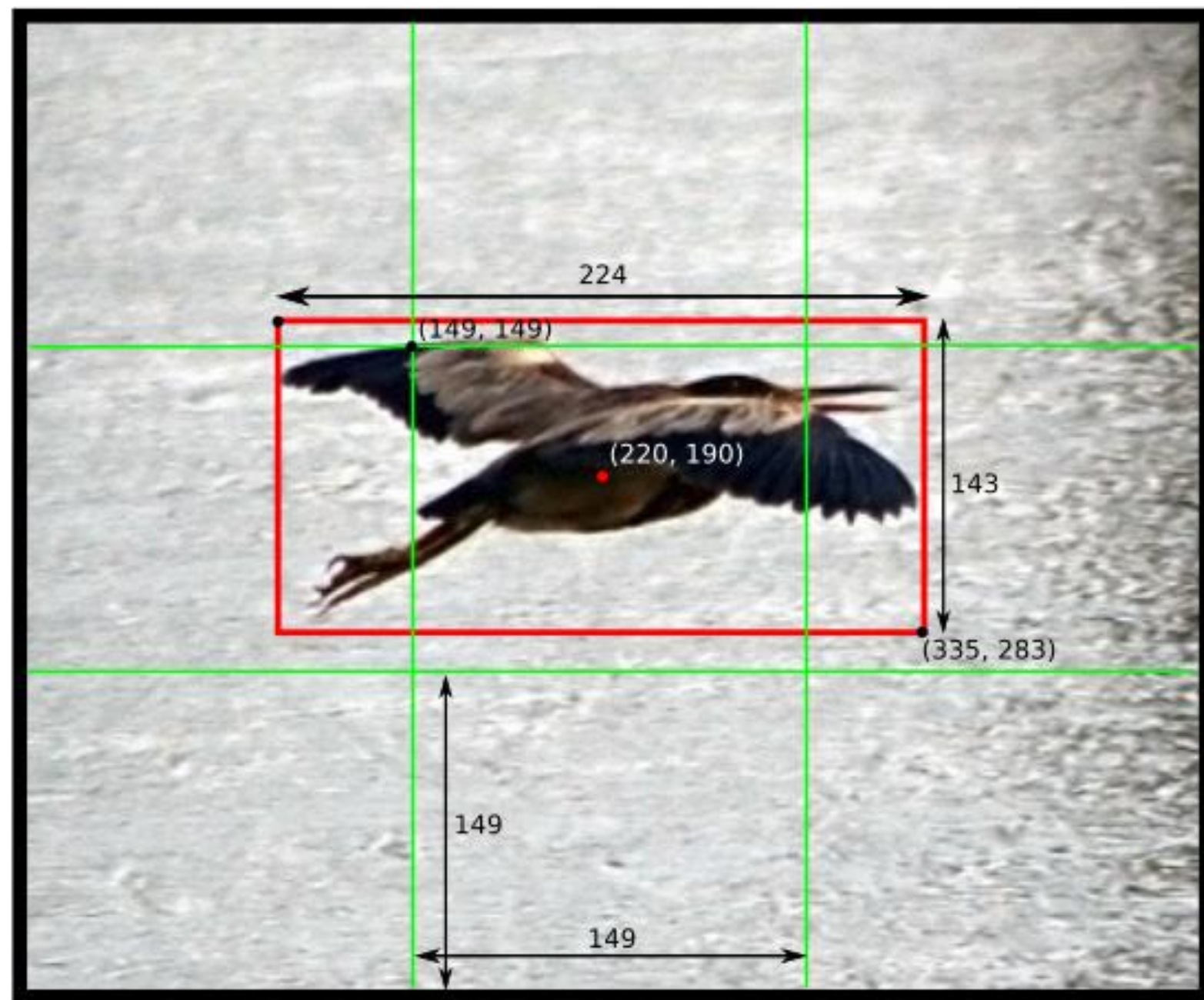
01. YOLO

02. YOLOv4

YOLO의 구조

Bounding Box Coordinates

(0, 0)



(447, 447)

$$x = (220 - 149) / 149 = 0.48$$

$$y = (190 - 149) / 149 = 0.28$$

$$w = 224 / 448 = 0.50$$

$$h = 143 / 448 = 0.32$$

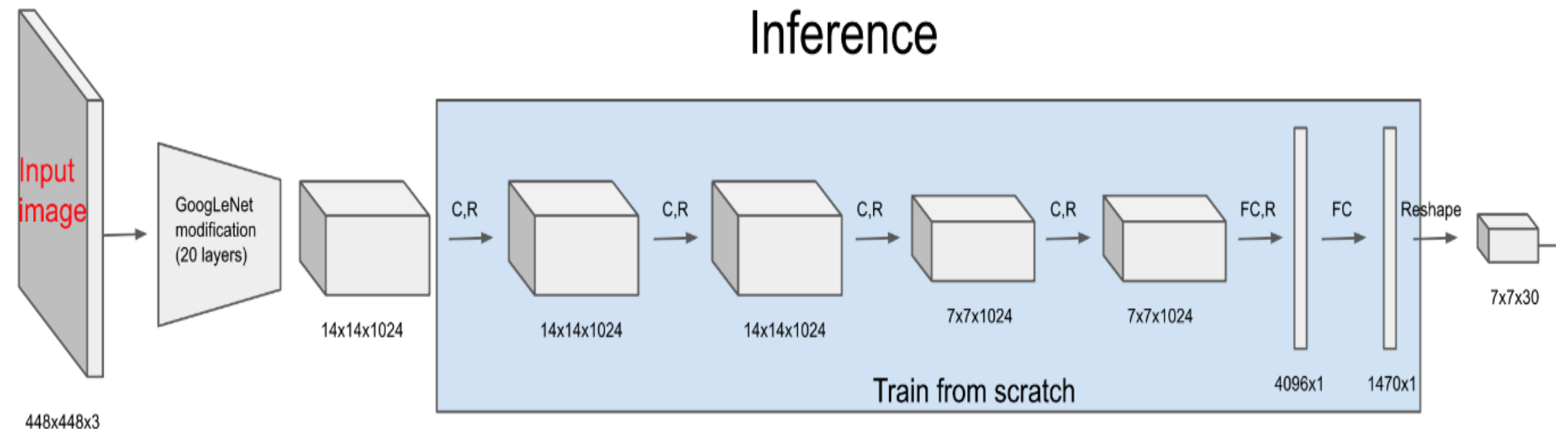
01

01. YOLO

02. YOLOv4

YOLO의 구조

Network Design



- 기본 Network Design은 GoogLeNet을 기반으로 함
- 24개의 Convolutional Layers + 2개의 Fully Connected Layer

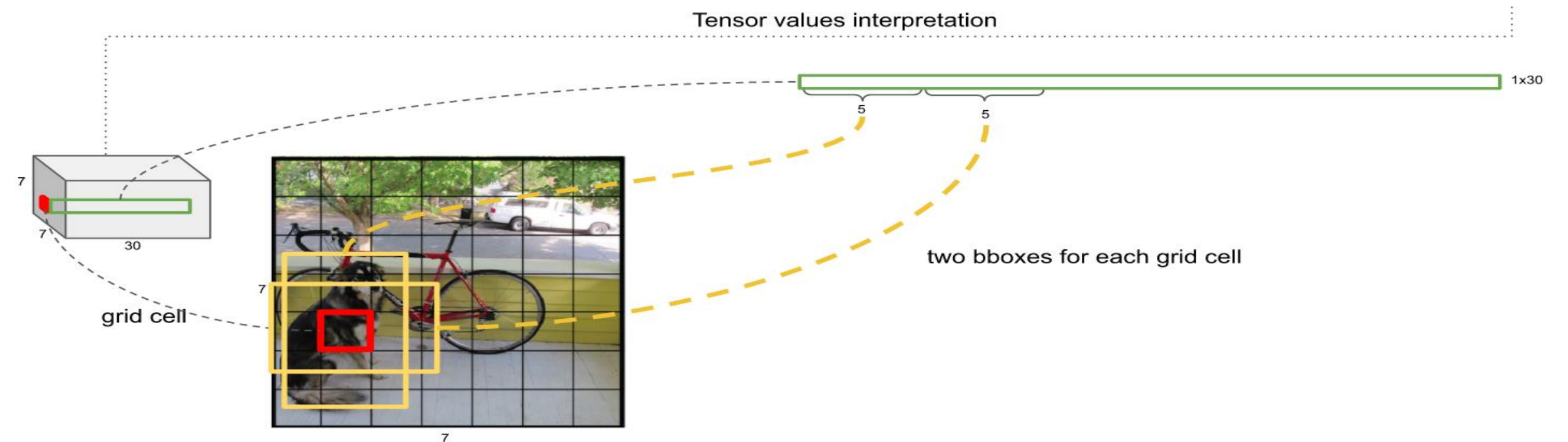
01

01. YOLO

02. YOLOv4

YOLO의 구조

Inference Process



- 최종적으로 7x7x30의 결과가 나옴
- 7x7 -> Grid Cell의 개수, 30개는 다음과 같음
 - 0~4는 첫 번째 bounding box의 5개 정보(x,y,w,d,cs)
 - 6~10 두 번째 bounding box의 5개 정보(x,y,w,d,cs)
 - 11~30 전체 20개 클래스에 대한 conditinal class probability

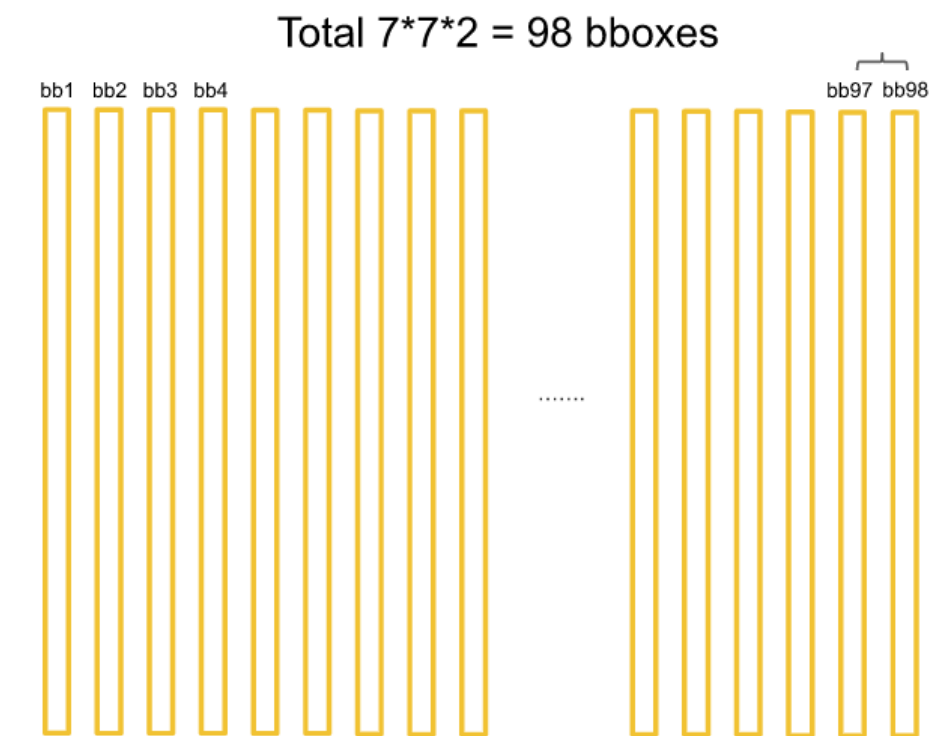
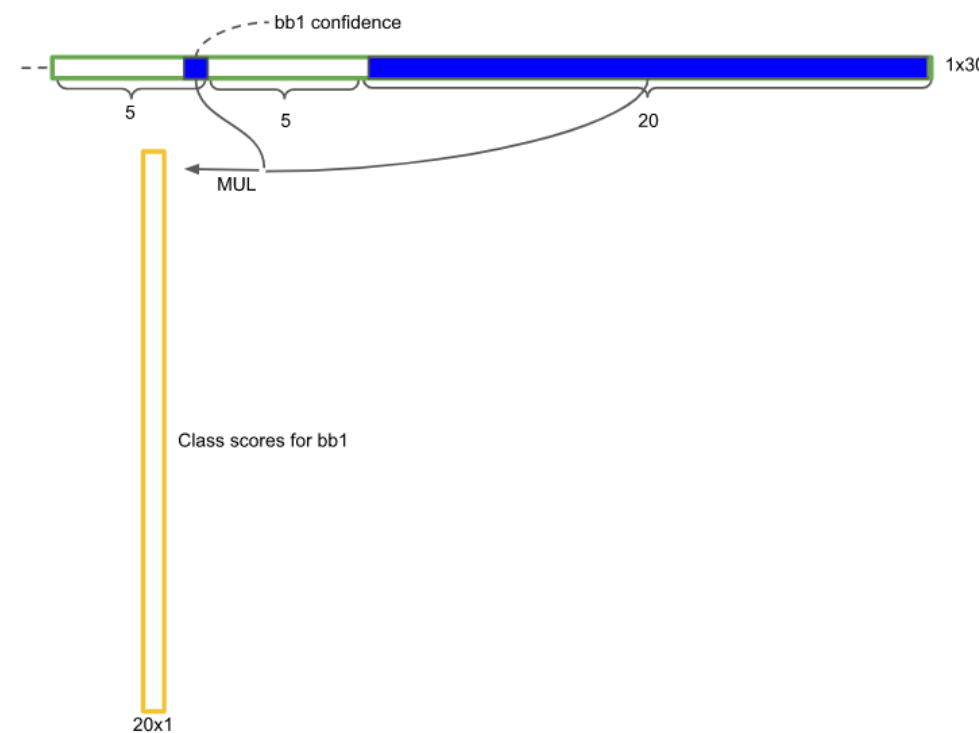
01

01. YOLO

02. YOLOv4

YOLO의 구조

Inference Process



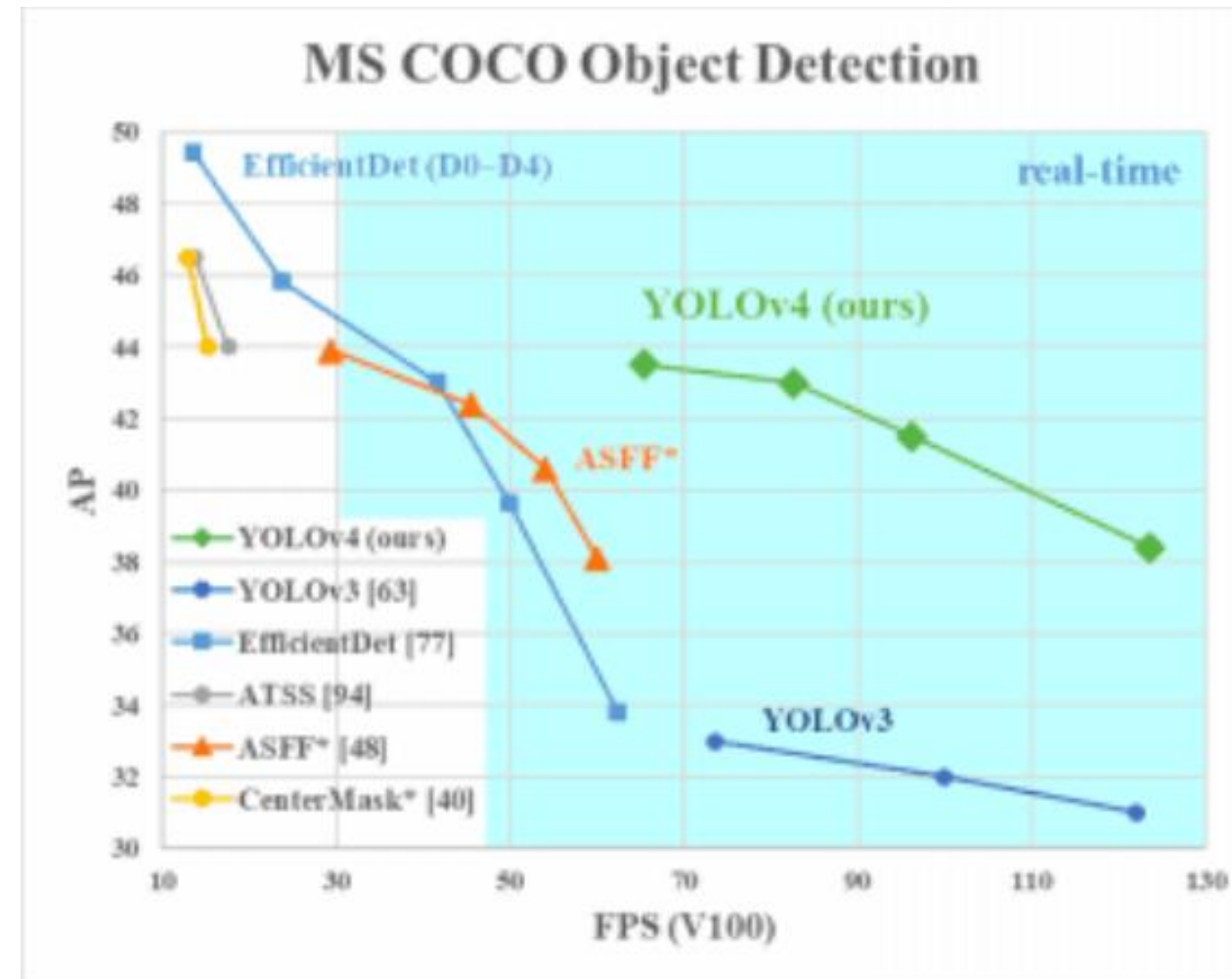
- 첫번째 bounding box의 Confidence Score를 20개의 Conditional Class Probability를 곱하면 class-specific confidence score가 나온다
- 총 $7 \times 7 \times 2 = 98$ 개의 class-specific confidence score가 나온다

01

01. YOLO

02. YOLOv4

Why YOLOv4?



- 가장 빠르고 정확한 가장 최신 object detection 기술
- 단 하나의 8~16GB GPU로만도 학습시킬 수 있음
- 작은 object를 타 모델들 보다 잘 검출함

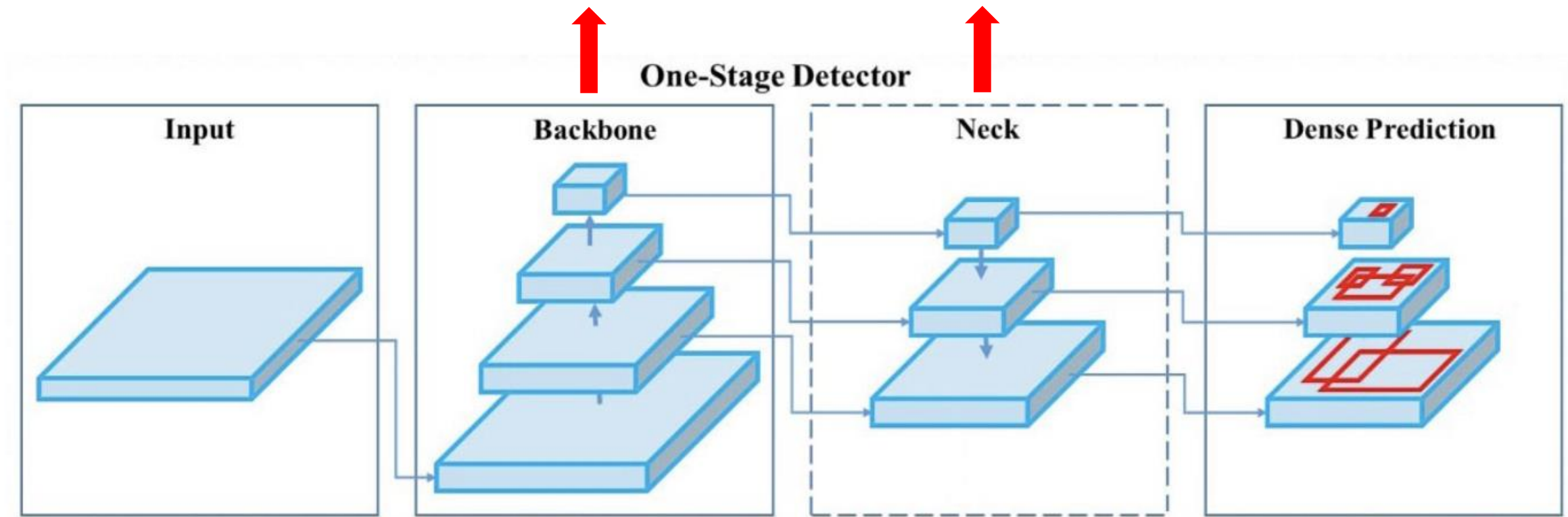
01

01. YOLO

02. YOLOv4

Model : YOLOv4

YOLOv4 =
YOLOv3 + CSPDarknet53 + SPP + PAN + BoF + BoS



01

01. YOLO

02. YOLOv4

Model : YOLOv4

Bag of Freebies – Training Phase

Data Augmentation

- Random Erase
- CutOut
- MixUp
- CutMix
- Style transfer GAN

Regularization

- DropOut
- DropPath
- Spatial DropOut
- DropBlock

Loss Function

- MSE
- IoU
- GloU
- CloU
- DloU

01

01. YOLO

02. YOLOv4

Model : YOLOv4

Bag of Specials – Inference Phase

Enhancement of Receptive Field

- Spatial Pyramid Pooling
- ASPP
- Receptive Field Block

Feature Integration

- Skip-connection
- Feature Pyramid
Network
- SFAM
- ASFF
- BiFPN

Activation Function

- ReLU
- Leaky ReLU
- Parametric ReLU
- ReLU6
- Swish
- Mish

참고 사이트

- <https://www.youtube.com/watch?v=Ae-p7QVOdbA>
- <https://www.youtube.com/watch?v=CXRlpsFpVUE>
- <https://ropiens.tistory.com/33>

Thank you

Itwill 12th LKYJ Team