Focal Loss for Dense Object Detection

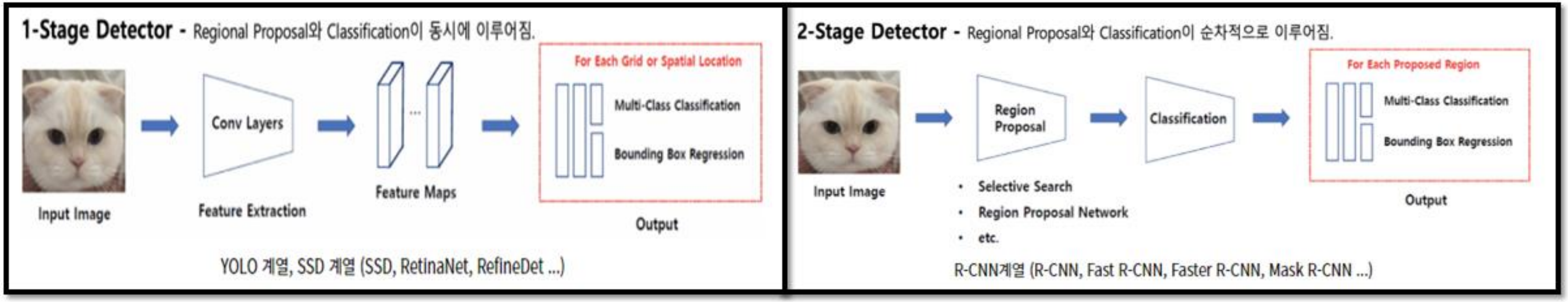
Seminar in 2023, Paper Review

Samsung Software Developer Community
Korea Vision & Robotics
Gi-Beom Kim
2023.06.10

https://openaccess.thecvf.com/content_iccv_2017/html/Lin_Focal_Loss_for_ICCV_2017_paper.html

Contents

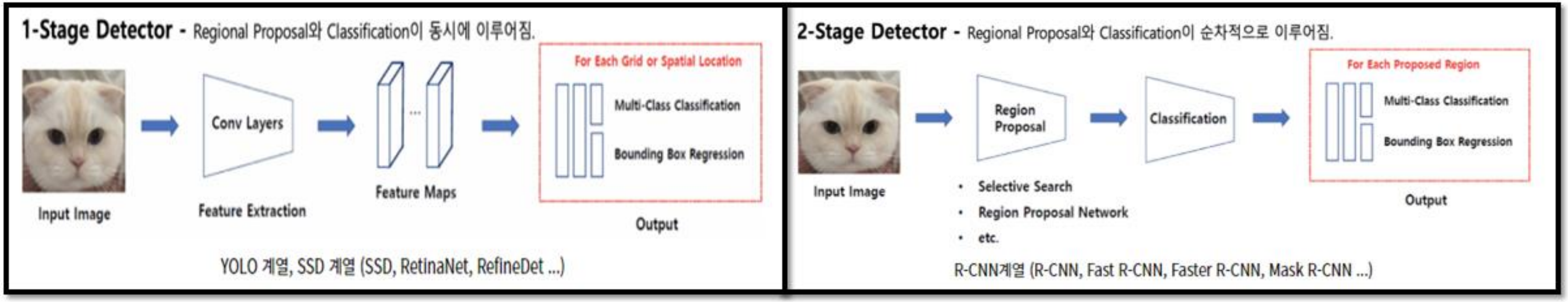
요약



1-Stage Detector
속도 빠름, 정확도 낮음

2-Stage Detector
속도 느림, 정확도 높음

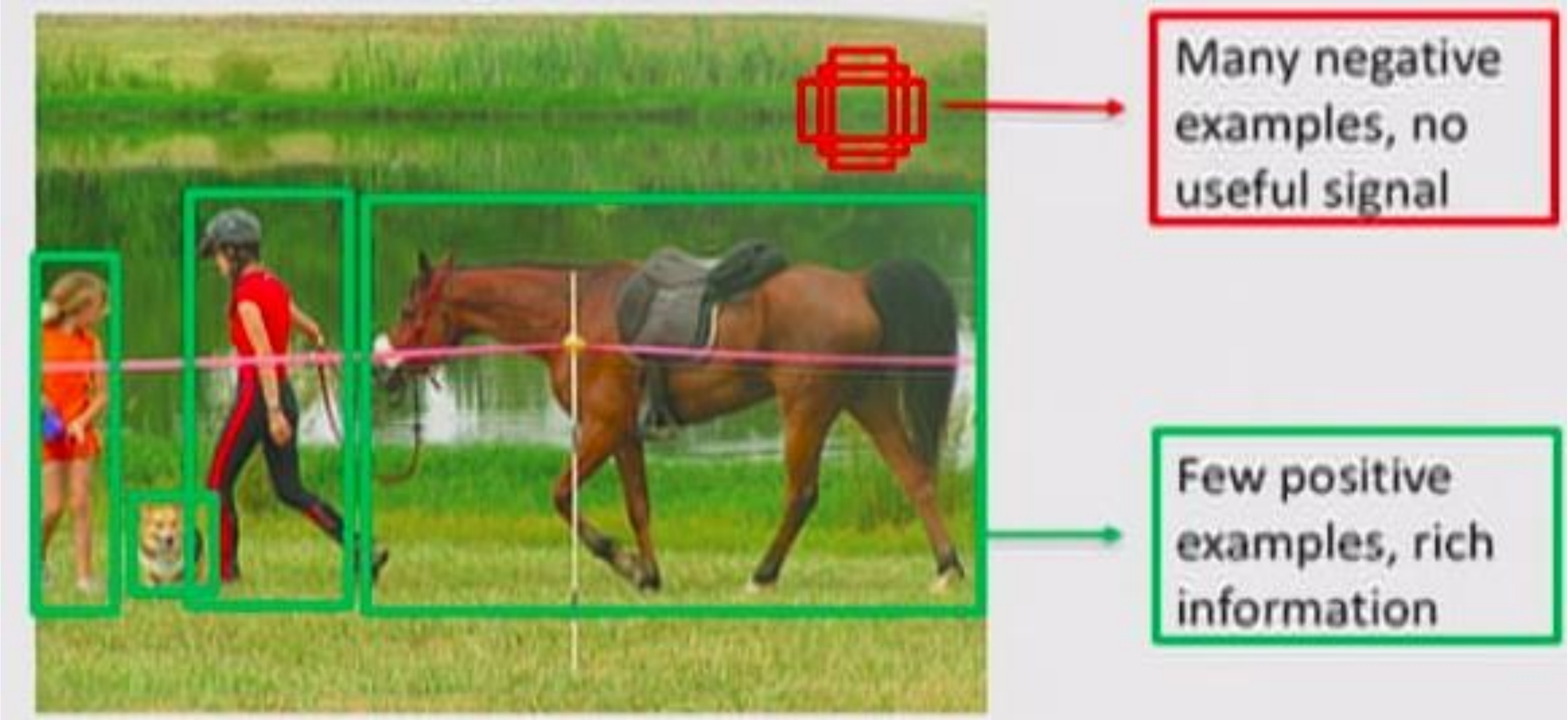
요약



1-Stage Detector
속도 빠름, 정확도 낮음
Localization, Classification 동시 진행

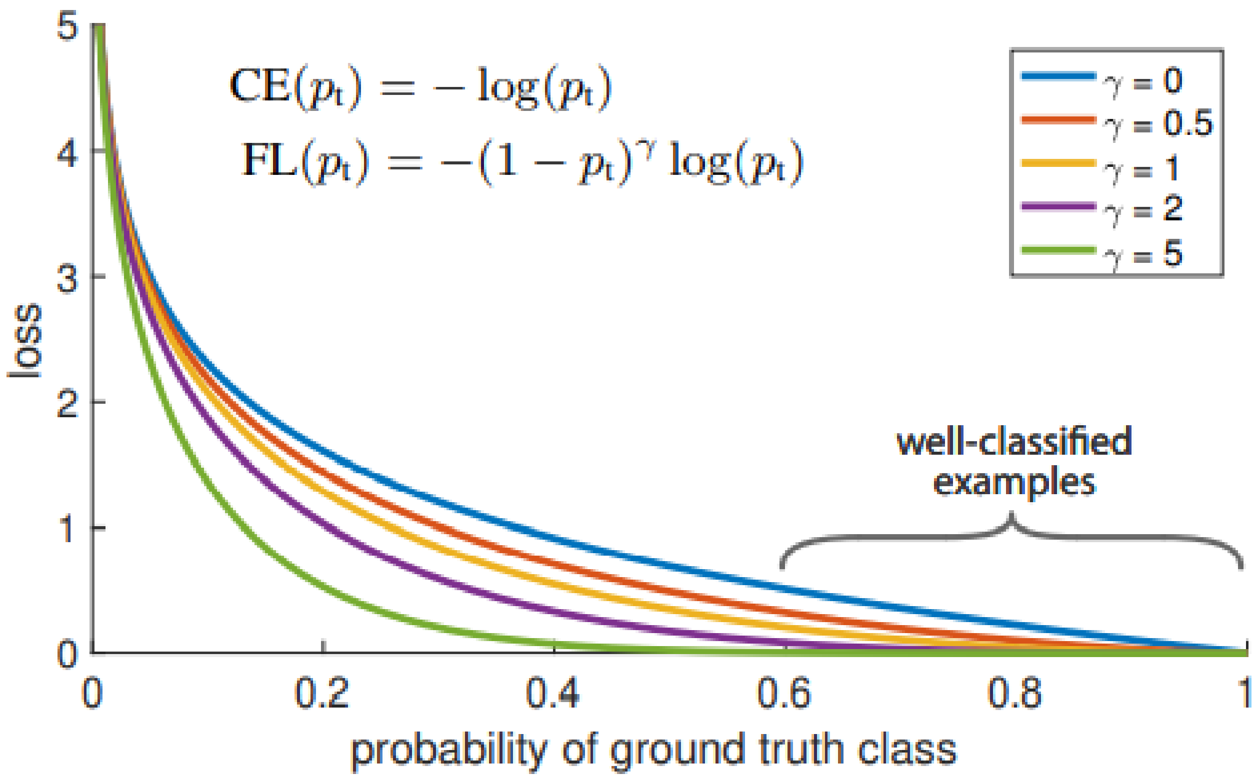
2-Stage Detector
속도 느림, 정확도 높음
1차: Localization, 2차: Classification

요약



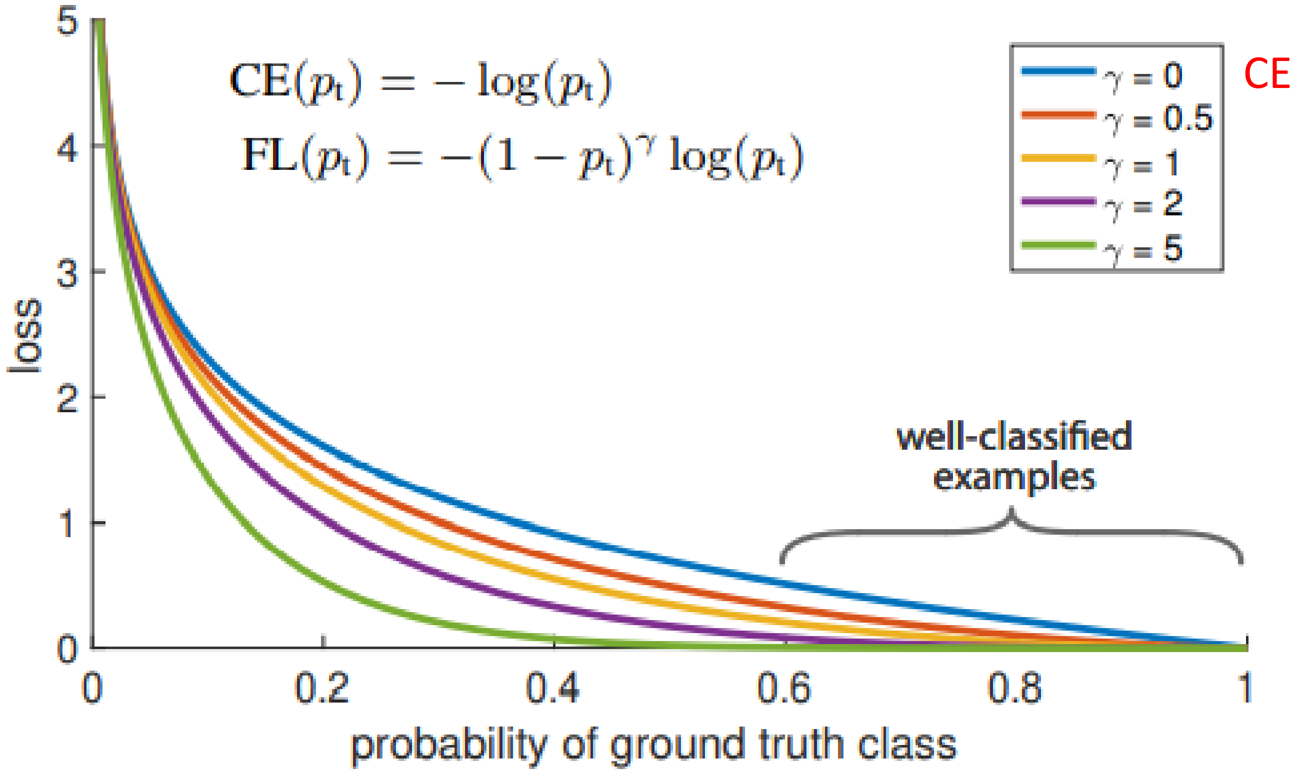
Focal loss는 쉬운 문제(초록색 박스)보다 어려운 문제(빨간 박스)에 가중치를 더 두어 문제를 해결하고자 함

요약



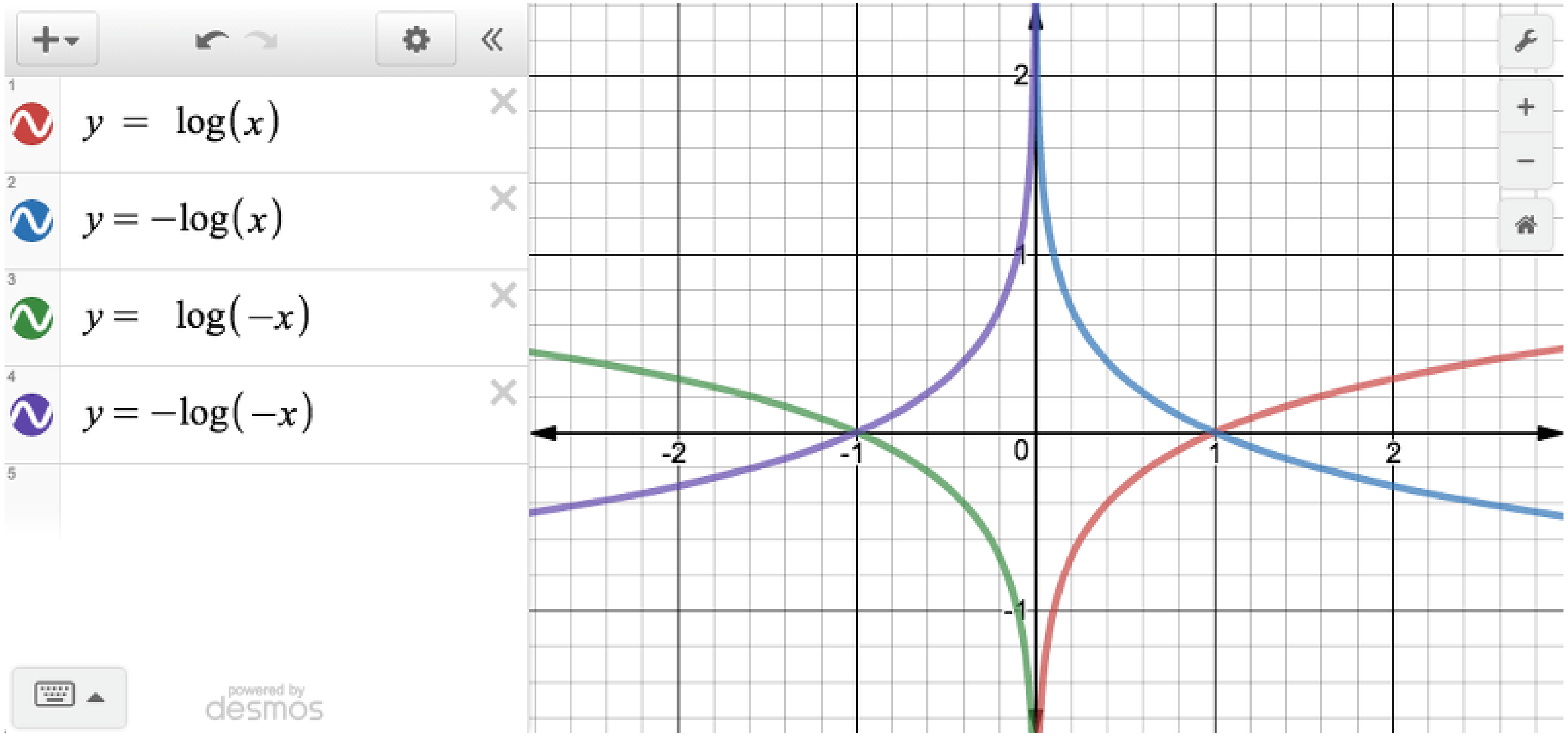
실제로 가중치(감마)를 크게 줄수록 loss가 낮아짐

Binary Cross Entropy Loss



$$\mathrm{CE}(p, y) = \begin{cases} -\log(p) & \text{if } y = 1 \\ -\log(1-p) & \text{otherwise.} \end{cases} \quad (1) \qquad p_t = \begin{cases} p & \text{if } y = 1 \\ 1-p & \text{otherwise,} \end{cases} \quad (2)$$

Binary Cross Entropy Loss



$$\mathrm{CE}(p, y) = \mathrm{CE}(p_t) = -\log(p_t).$$

Cross Entropy Loss

$$H(P,Q) = -\mathbb{E}_{X\sim P}[\log Q(x)] = -\sum P(x)\log Q(x)$$

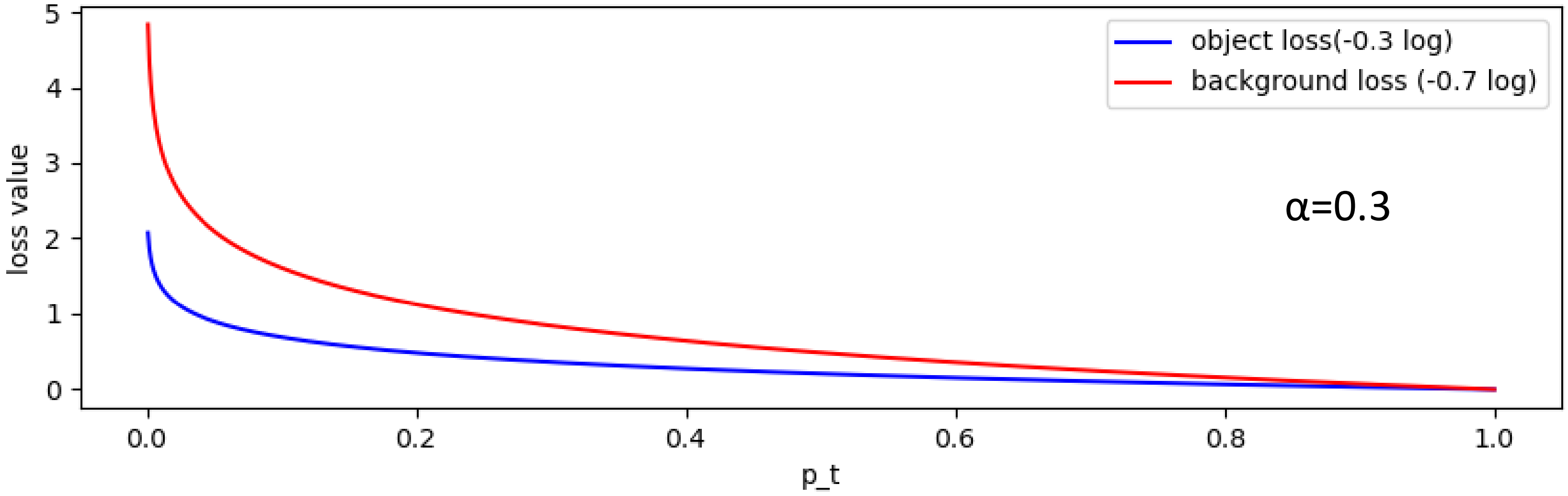
$$\text{softmax}(\text{input}) = \begin{bmatrix} 0.2 \\ 0.7 \\ 0.1 \end{bmatrix}$$

$$Q(X=c_1)=0.2$$
$$Q(X=c_2)=0.7$$
$$Q(X=c_3)=0.1$$

$$P(X=c_1)=0$$
$$P(X=c_2)=1$$
$$P(X=c_3)=0$$

$$\begin{aligned} H(P,Q) &= -\sum P(x)\log Q(x) \\ &= -(0\cdot\log 0.2 + 1\cdot\log 0.7 + 0\cdot\log 0.1) \\ &= -\log 0.7 \approx 0.357 \end{aligned}$$

Balanced Binary Cross Entropy Loss



$$\mathrm{CE}(p_t) = -\alpha_t \log(p_t).$$

- class 불균형을 해결하기 위해 임의의 가중치 **α** 적용
- Focal Loss의 기반이 되는 수식
- Positive example(object): α
- Negative example(background): 1 – α

Focal Loss

$$\mathrm{FL}(p_t) = -(1 - p_t)^{\gamma} \log(p_t). \qquad (4)$$

- Balanced CE 장점: 단일 분류(긍정/부정)에 대해 가중치 적용 가능
- Balanced CE 단점: example들에 대한 가중치 적용 불가능

- Easy example에 대해 가중치를 줄여 loss를 낮게 하고
- Hard example에 대해 가중치를 높여 loss를 높게 학습함

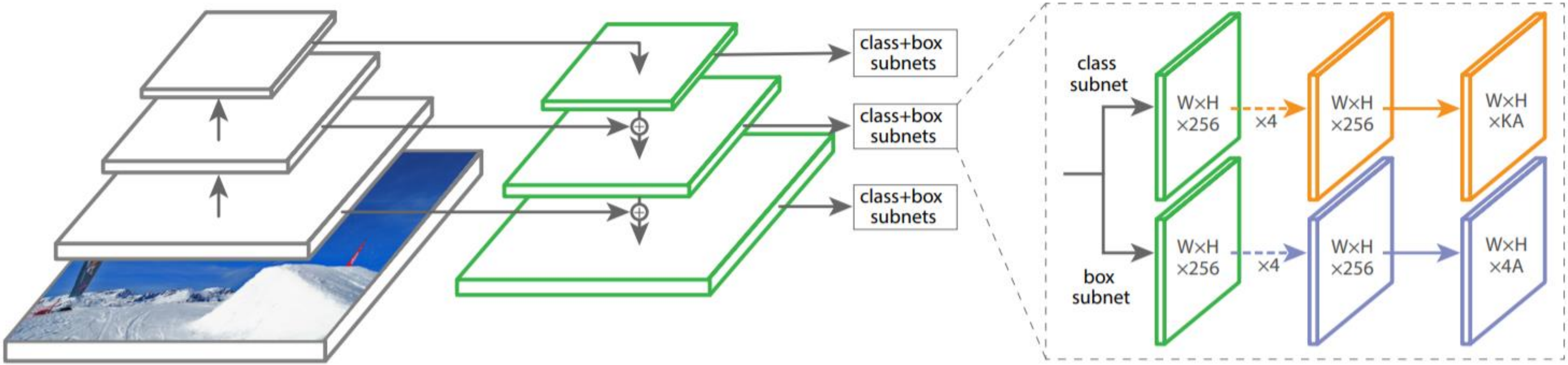
- P_t가 작으면 맞추기 어려운 example == hard example
- P_t가 작으면 loss가 커짐
- P_t가 크면 loss가 작아짐
- 감마가 커질수록 (1-p_t)의 영향력이 커짐 (논문에서는 감마=2가 베스트)

Balanced Focal Loss

$$\mathrm{FL}(p_t) = -\alpha_t (1 - p_t)^\gamma \log(p_t). \qquad (5)$$

- Balanced Cross Entropy와 같은 알파 적용
- Focal loss보다 좋은 성능을 보여줌

RetianNet



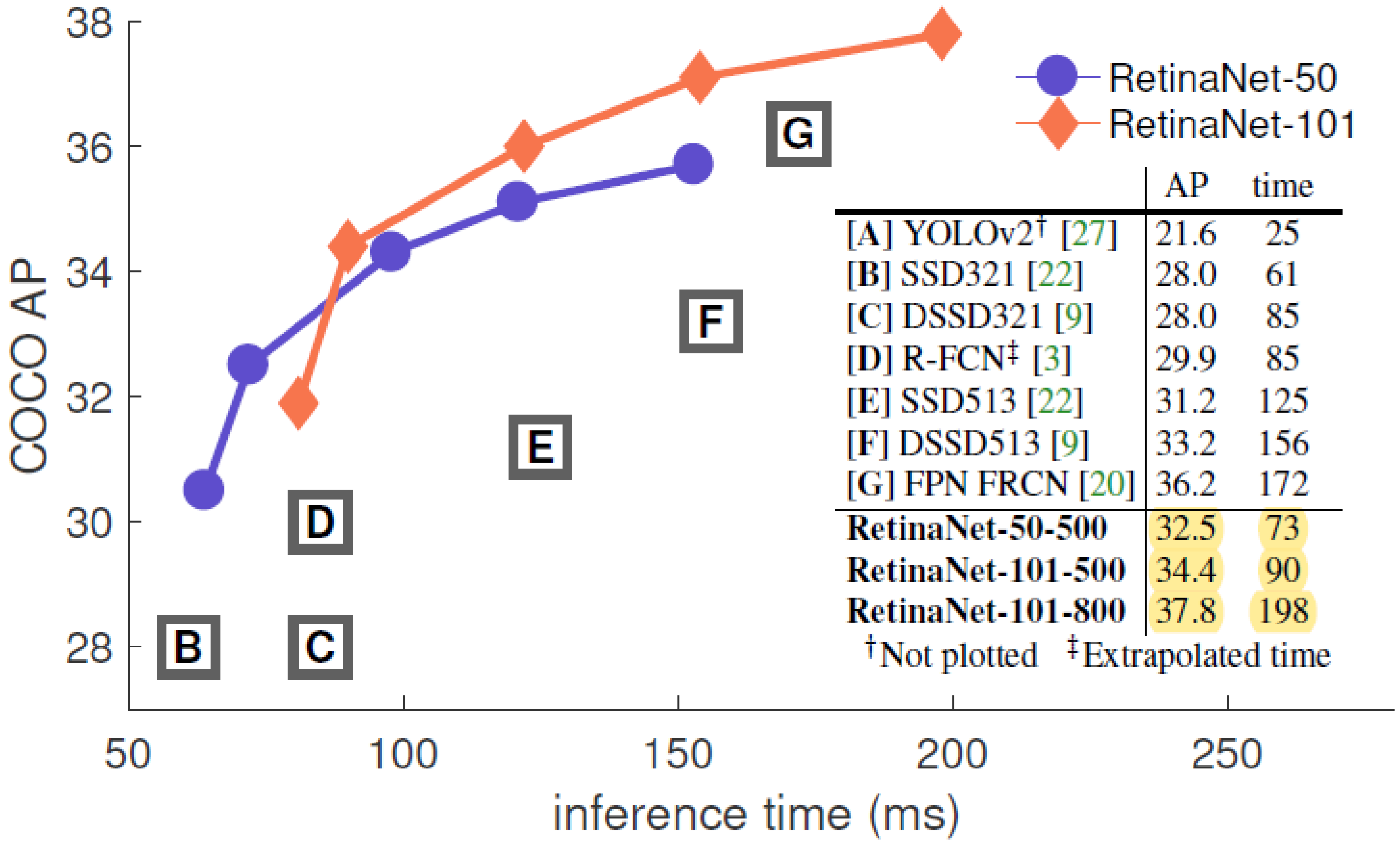
(a) ResNet (b) feature pyramid net (c) class subnet (top) (d) box subnet (bottom)

Object를 분류

Box를 제대로 그렸는지 연산

- One Stage Detection을 위한 아키텍처
- ResNet, FPN 네트워크 채용
- 두 가지 subnet 보유

Results


38
36
34
32
30
28
COCO AP
50
100
150
200
250
inference time (ms)
RetinaNet-50
RetinaNet-101
B
C
D
E
F
G
AP time
[A] YOLOv2† [27] 21.6 25
[B] SSD321 [22] 28.0 61
[C] DSSD321 [9] 28.0 85
[D] R-FCN‡ [3] 29.9 85
[E] SSD513 [22] 31.2 125
[F] DSSD513 [9] 33.2 156
[G] FPN FRCN [20] 36.2 172
RetinaNet-50-500 32.5 73
RetinaNet-101-500 34.4 90
RetinaNet-101-800 37.8 198
†Not plotted ‡Extrapolated time

Results

	backbone	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
Two-stage methods							
Faster R-CNN+++ [16]	ResNet-101-C4	34.9	55.7	37.4	15.6	38.7	50.9
Faster R-CNN w FPN [20]	ResNet-101-FPN	36.2	59.1	39.0	18.2	39.0	48.2
Faster R-CNN by G-RMI [17]	Inception-ResNet-v2 [34]	34.7	55.5	36.7	13.5	38.1	52.0
Faster R-CNN w TDM [32]	Inception-ResNet-v2-TDM	36.8	57.7	39.2	16.2	39.8	**52.1**
One-stage methods							
YOLOv2 [27]	DarkNet-19 [27]	21.6	44.0	19.2	5.0	22.4	35.5
SSD513 [22, 9]	ResNet-101-SSD	31.2	50.4	33.3	10.2	34.5	49.8
DSSD513 [9]	ResNet-101-DSSD	33.2	53.3	35.2	13.0	35.4	51.1
RetinaNet (ours)	ResNet-101-FPN	39.1	59.1	42.3	21.8	42.7	50.2
RetinaNet (ours)	ResNeXt-101-FPN	**40.8**	**61.1**	**44.1**	**24.1**	**44.2**	51.2

감사합니다.

Seminar in 2023, Paper Review

Samsung Software Developer Community
Korea Vision & Robotics
Gi-Beom Kim
2023.06.10

https://openaccess.thecvf.com/content_iccv_2017/html/Lin_Focal_Loss_for_ICCV_2017_paper.html