Focal Loss for Dense Object Detection
ICCV 2017

Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár

总结

- Focal Loss：为解决正负样本数目不均衡而提出的一种损失函数
  - 在物体检测中，负样本通常比正样本多很多

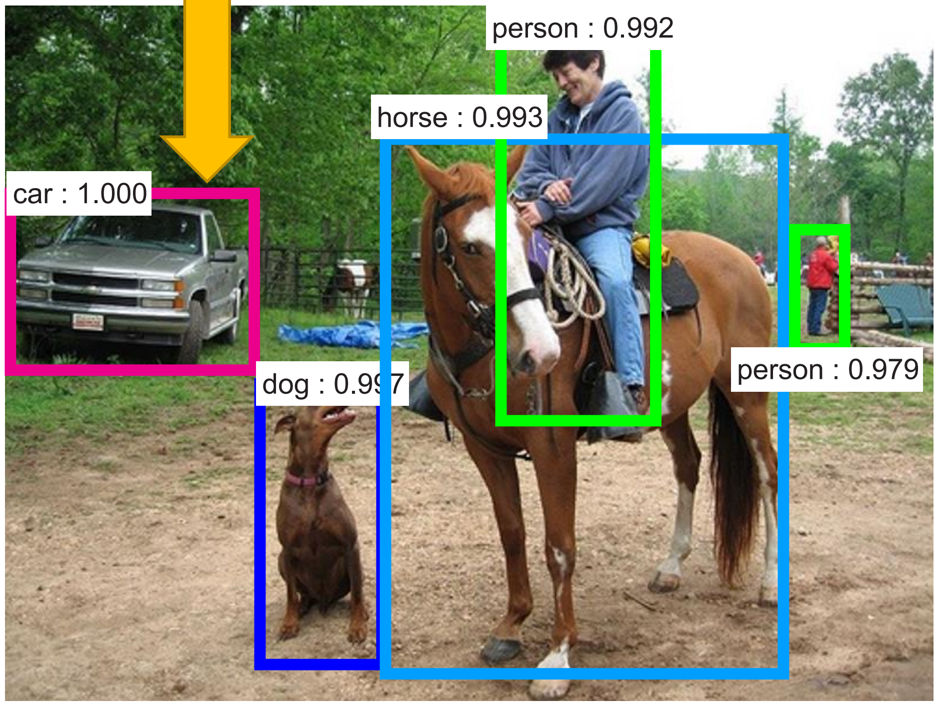
▶ 物体检测

Localization

Where?

Recognition

What?



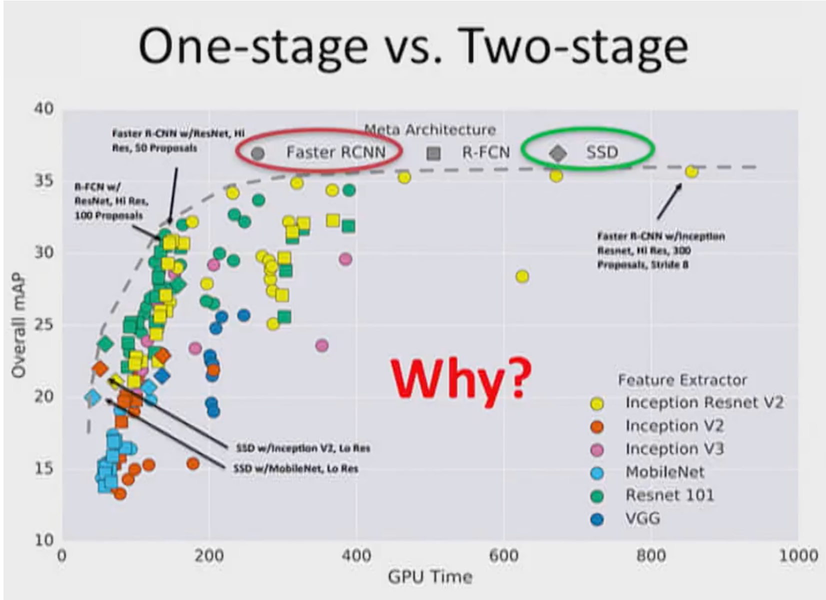
训练样本

- 正样本：标注
- 负样本：从背景区域随机采样

- 因此：负样本数目比正样本多很多

▶ 主流物体检测框架

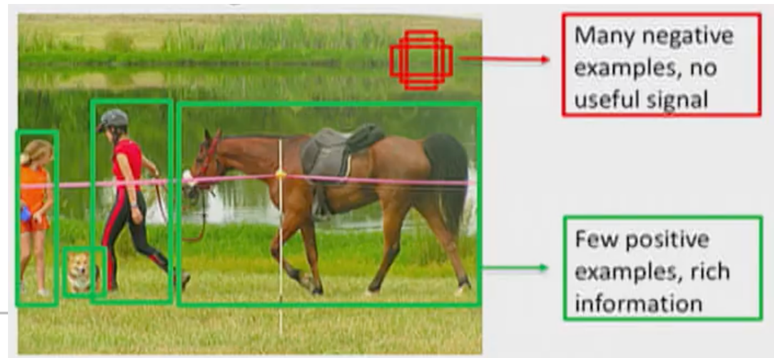
- One-Stage Detector： SSD/YOLO
  - 在每个位置输出分类结果/稠密proposal
  - 速度快但精度低
- Two-Stage Detector： Faster-RCNN
  - 先提少量proposal，再分类
  - 速度慢但精度高



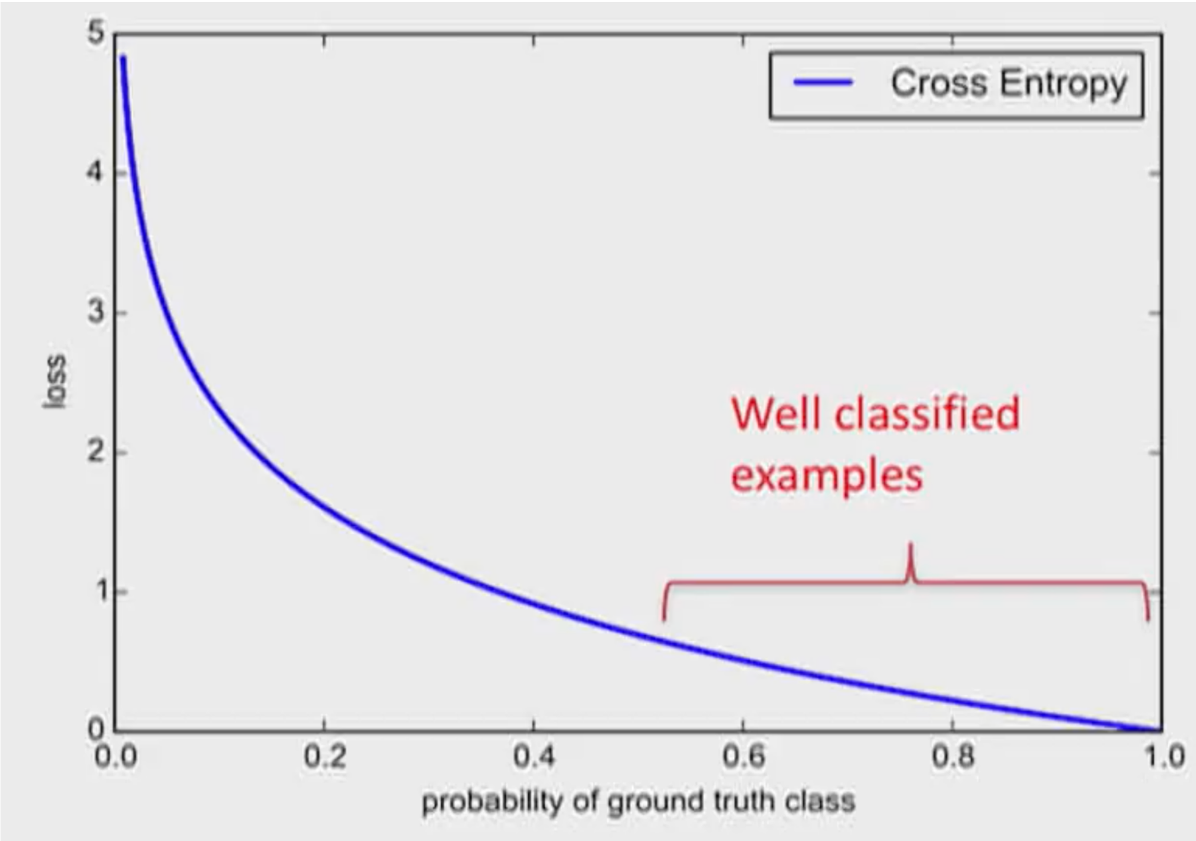
Speed/accuracy trade-offs for modern convolutional object detectors, Huang et al., CVPR 2017

▶为什么One-Stage Detector精度不高

- 更稠密的box： ~ 100K
- 稠密proposal中，前景和背景的极度不平衡
  - 如在PASCAL VOC数据集中，每张图片上标注的目标：几个
  - YOLO V2最后一层的输出：13×13×513×13×5，845845个候选目标！
  - 大量（简单易区分）的负样本在loss中占据了很大比重，使得有用的loss不能回传回来



▶交叉熵损失/loglosss



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

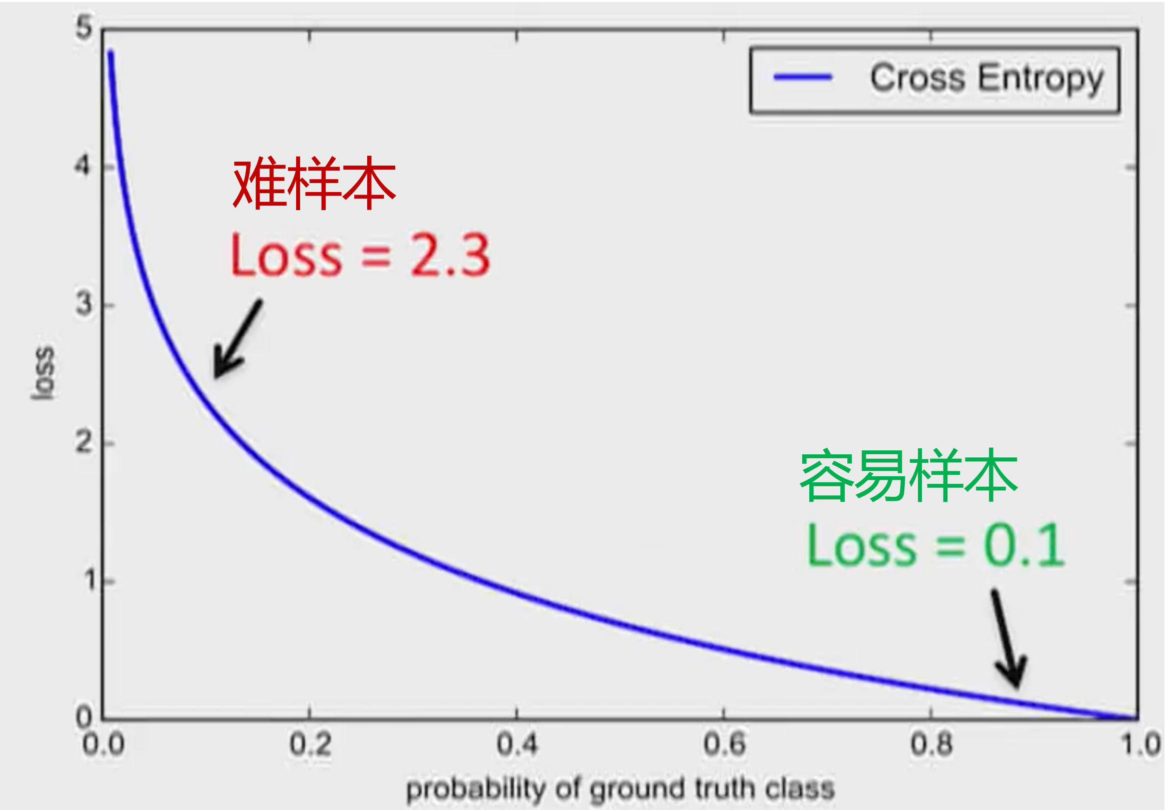
定义Probability of ground truth class为

$$p_{\mathrm{t}} = \begin{cases} p & \text{if } y = 1 \\ 1-p & \text{otherwise,} \end{cases}$$

则

$$\mathrm{CE}(p, y) = \mathrm{CE}(p_{\mathrm{t}}) = -\log(p_{\mathrm{t}})$$

▶交叉熵损失/loglosss



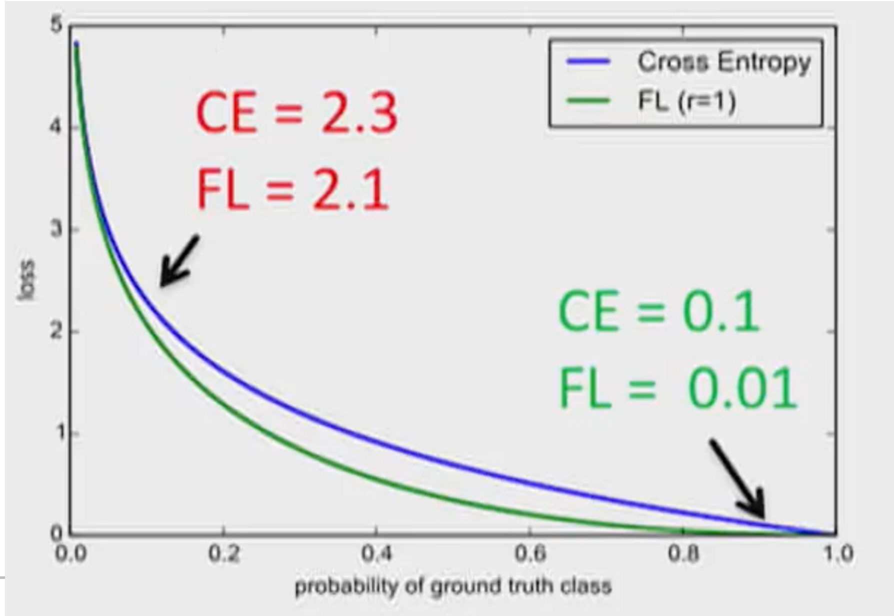
100000容易样本：100难样本

来自容易样本的loss为来自难样本的40倍

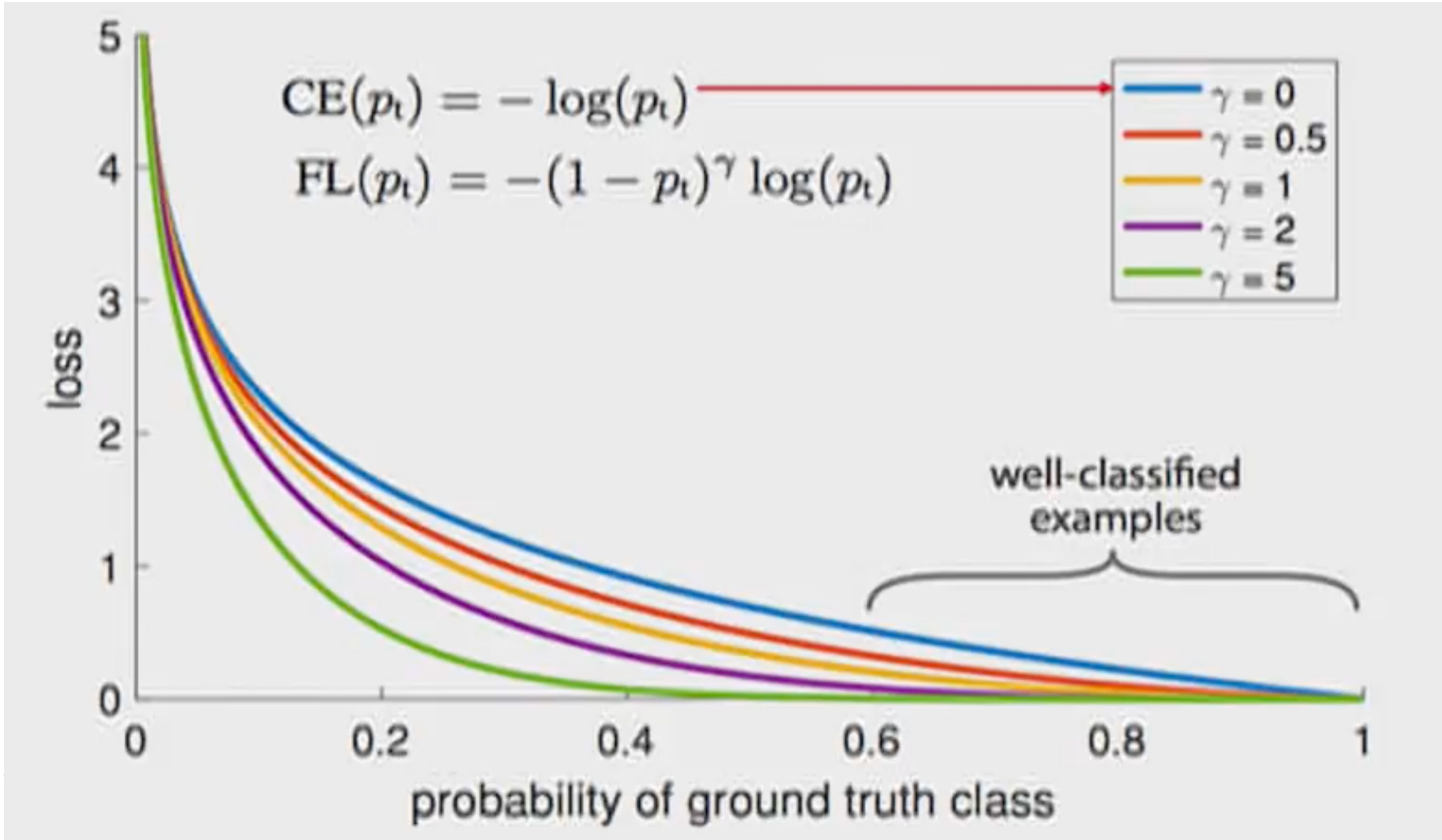
focal loss

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

$$\mathrm{FL}(p_t) = -\boxed{(1-p_t)^\gamma}\log(p_t)$$



focal loss


5
4
3
2
1
0
loss
$CE(p_t) = -\log(p_t)$
$FL(p_t) = -(1-p_t)^{\gamma}\log(p_t)$
$\gamma = 0$
$\gamma = 0.5$
$\gamma = 1$
$\gamma = 2$
$\gamma = 5$
well-classified
examples
0
0.2
0.4
0.6
0.8
1
probability of ground truth class

α-balanced variant of the focal loss

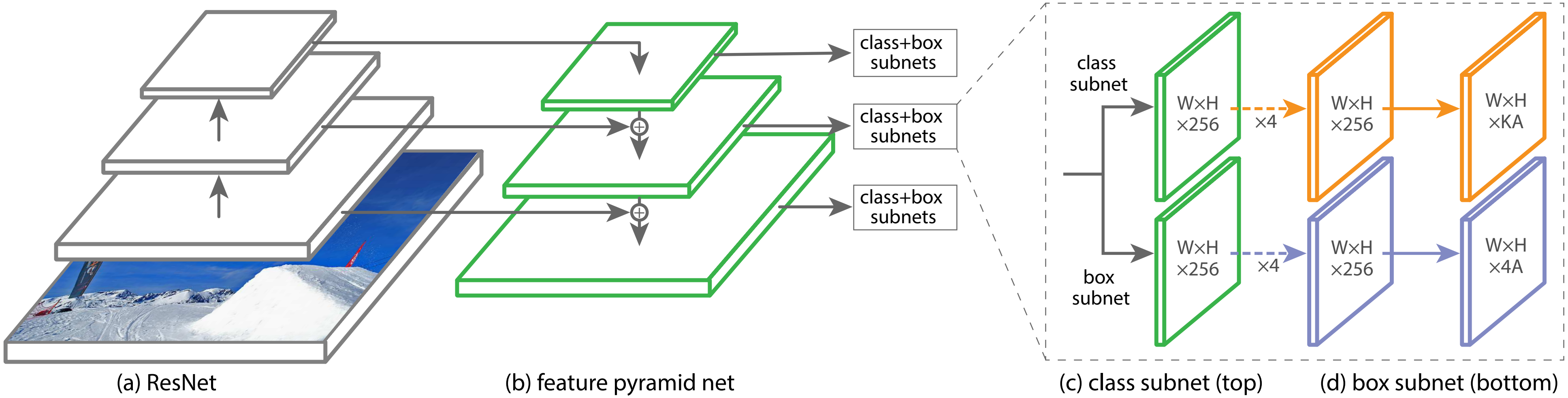
$$\mathrm{CE}(p_{\mathrm{t}}) = -\alpha_{\mathrm{t}} \log(p_{\mathrm{t}})$$

$$\mathrm{FL}(p_{\mathrm{t}}) = -\alpha_{\mathrm{t}}(1 - p_{\mathrm{t}})^{\gamma} \log(p_{\mathrm{t}})$$

- γ：对难样本施加更多权重
- α：不同类别的样本数目不均衡

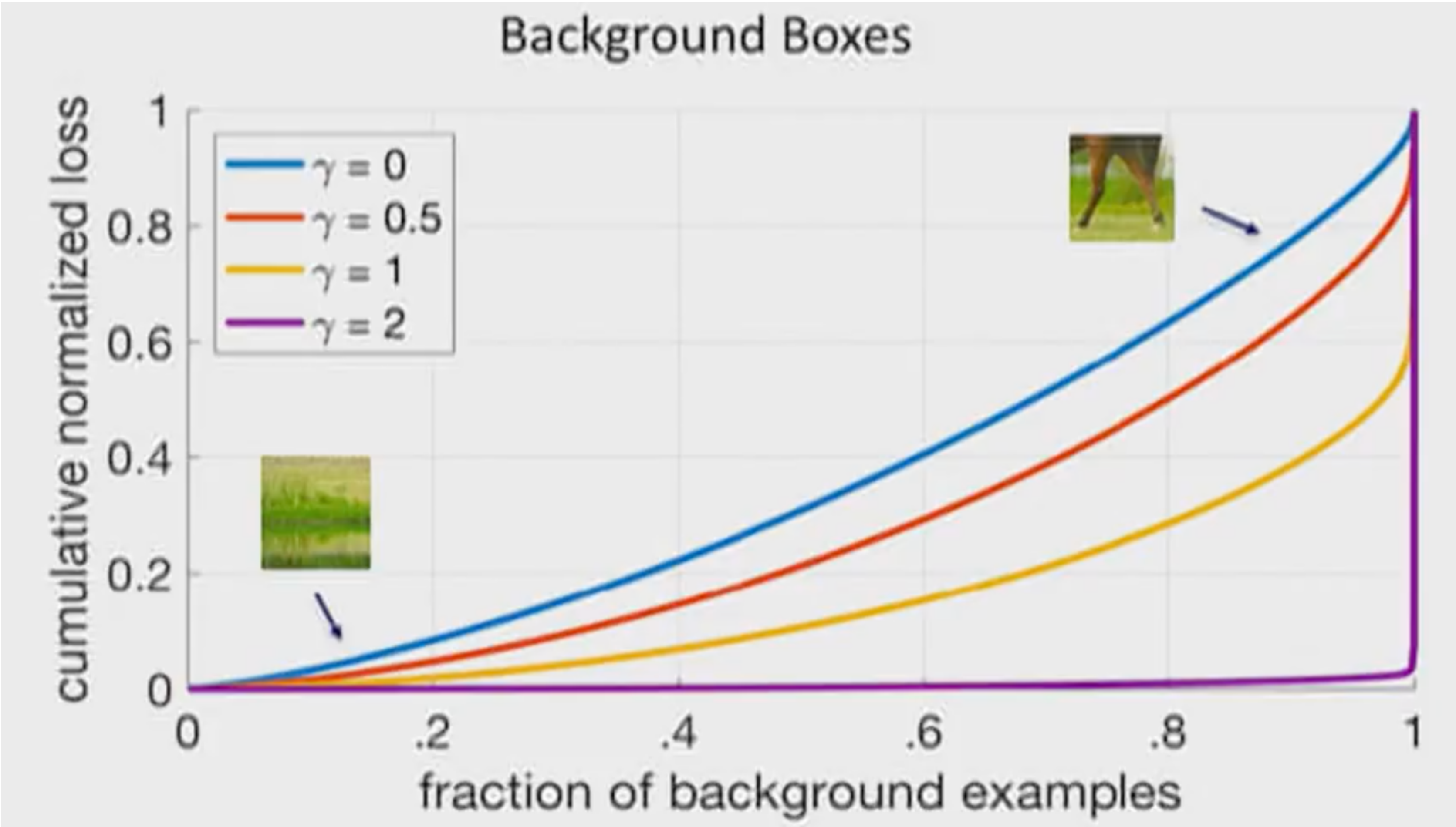
▶ 物体检测

- 检测网络框架：RetinaNet
- 损失： focal loss



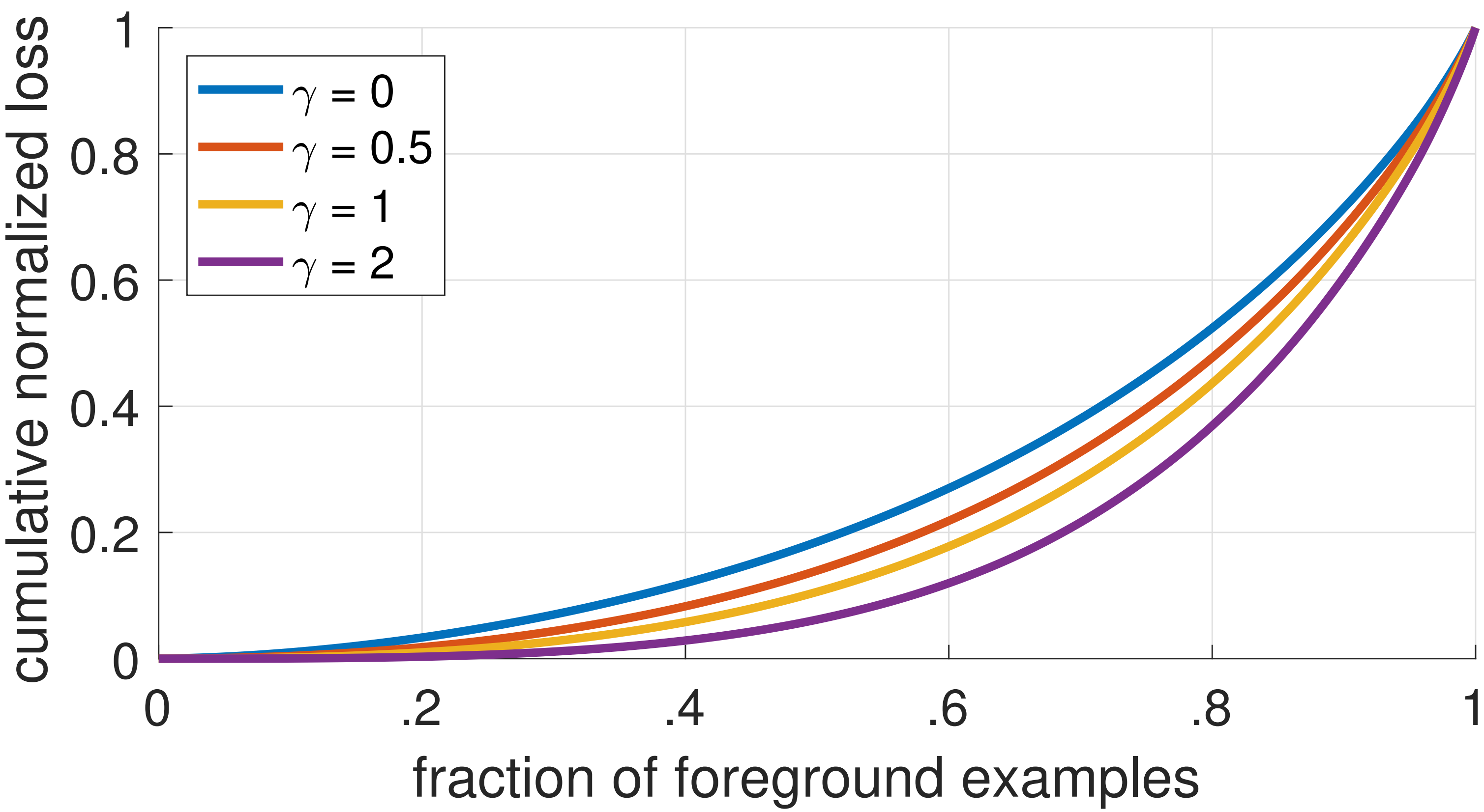
模型参数初始化：模型初始化之后，模型输出稀有类别的概率变小（如0.01）

Focal loss在背景样本上的累积分布函数



对背景中的难样本施加大的权重

Focal loss在前景样本上的累积分布函数



对前景样本本影响不大

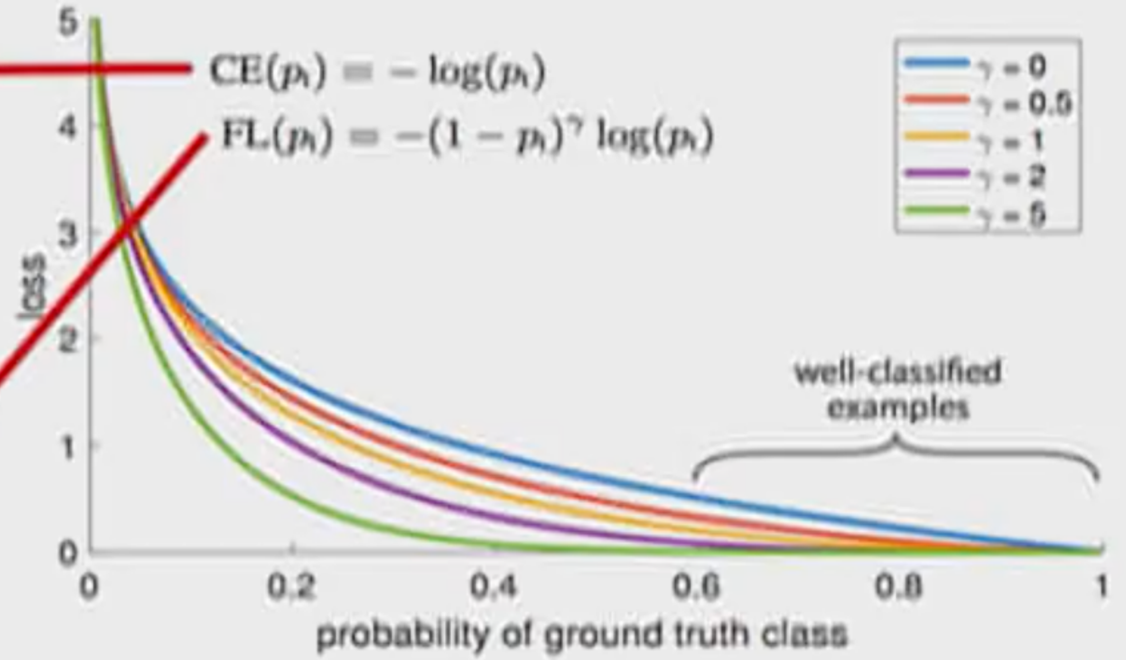
►物体检测结果——与交叉熵损失比较

CSDN 不止于代码

γ	α	AP	AP_{50}	AP_{75}
0	.75	31.1	49.4	33.0
0.1	.75	31.4	49.9	33.1
0.2	.75	31.9	50.7	33.4
0.5	.50	32.9	51.7	35.2
1.0	.25	33.7	52.0	36.2
2.0	.25	**34.0**	**52.5**	**36.5**
5.0	.25	32.2	49.6	34.8

(b) **Varying γ for FL** (w. optimal α)

(ResNet-50-FPN 600px input image)



AP提升2.9

▶物体检测结果——与OHEM比较

method	batch size	nms thr	AP
OHEM	128	.7	31.1
OHEM	256	.7	31.8
OHEM	512	.7	30.6
OHEM	128	.5	32.8
OHEM	256	.5	31.0
OHEM	512	.5	27.6
OHEM 1:3	128	.5	31.1
OHEM 1:3	256	.5	28.3
OHEM 1:3	512	.5	24.0
FL	n/a	n/a	**36.0**

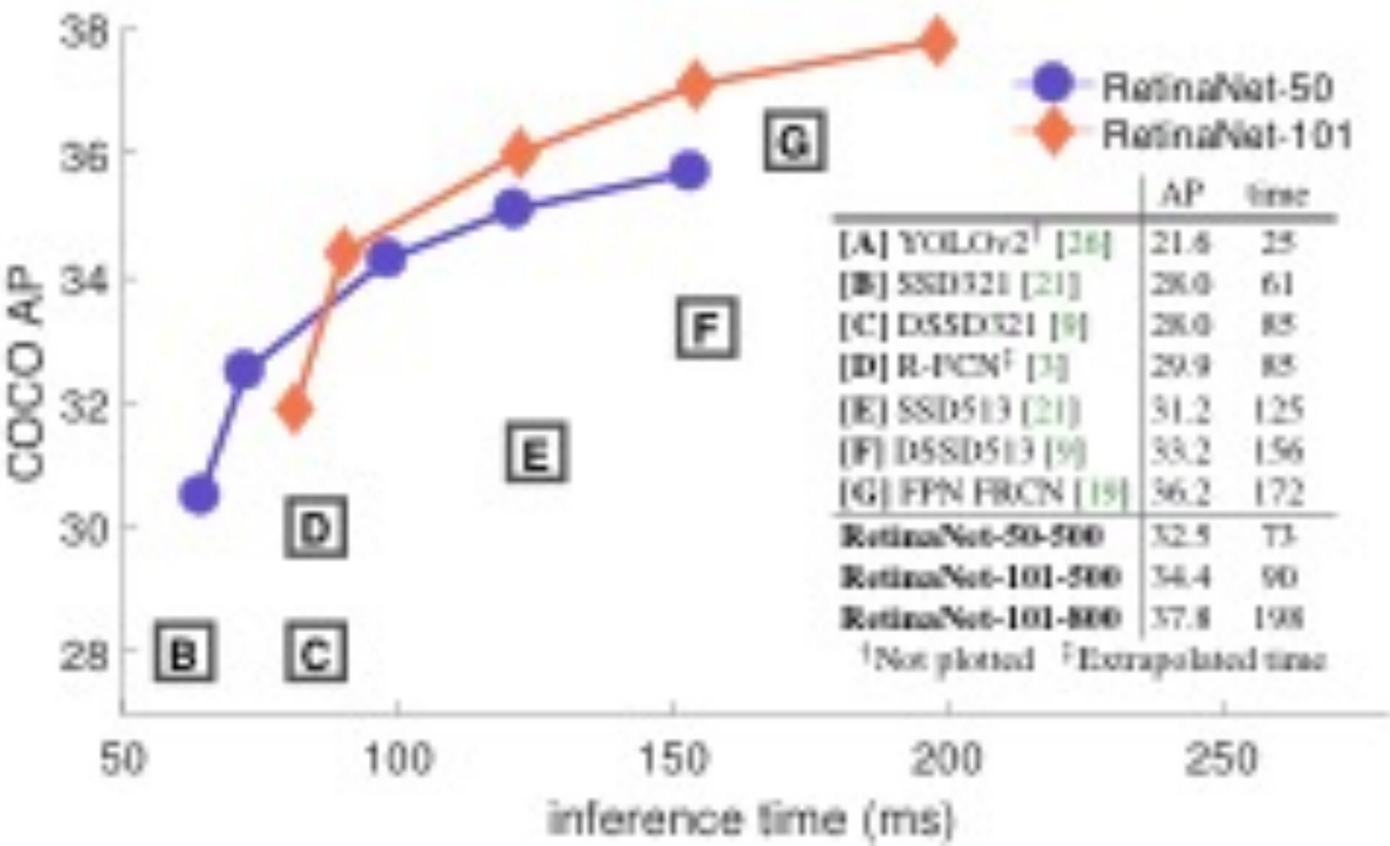
→ Best OHEM

→ Best Focal Loss

Online Hard Example Mining, *Shrivastava et al., 2016*

AP提升3.2

▶物体检测结果——与sate-of-art结果比较



精度达到与Two Stage最好结果

速度更快

▶ 在猫狗分类任务上实验Focal Loss

- Focal Loss似乎对样本不均衡的分类问题表现并不比交叉熵损失好
  - 检测中难样本可能才是关键
- https://shaoanlu.wordpress.com/2017/08/16/applying-focal-loss-on-cats-vs-dogs-classification-task/#more-1302
- https://github.com/shaoanlu/expriment-with-focal-loss/tree/master

参考链接

- http://blog.csdn.net/u014380165/article/details/77019084
- https://xmfbit.github.io/2017/08/14/focal-loss-paper/
- https://github.com/facebookresearch/Detectron