**Noisy label detection approach**

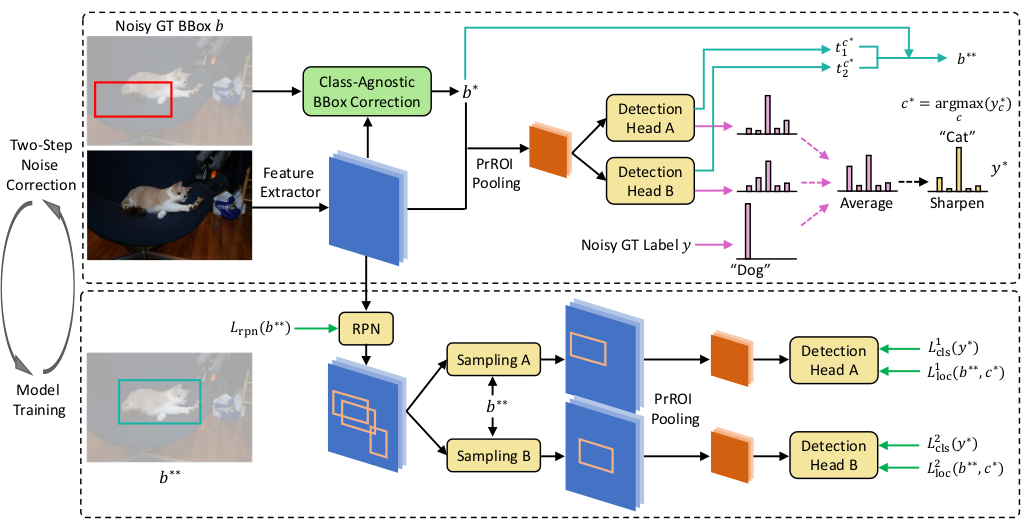
Based on object detection(not image classification)

**Focus on Noisy Label Detection**

# Two-Step Noise Correction Procedure:

Reference: <https://arxiv.org/pdf/2003.01285.pdf>

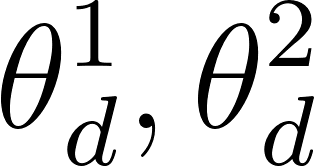
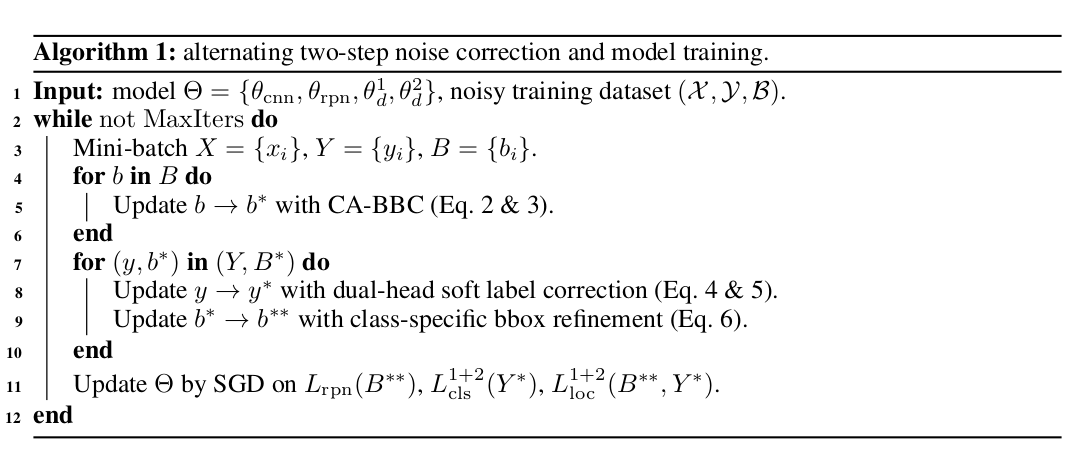
Closed source



Pipeline:

1. Warm-up: Train the detector with the origin whole dataset

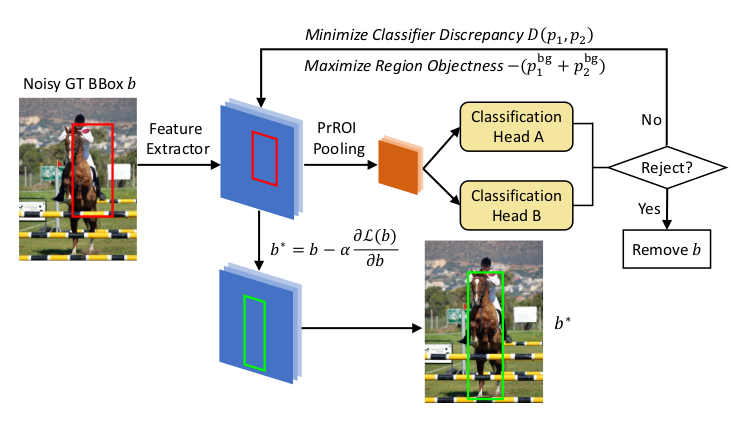
2. Noise samples correlation

Using Faster-RCNN as object detector, which has a backbone feature extractor parameterized by[](https://www.codecogs.com/eqnedit.php?latex=%20%5Ctheta_%7Bcnn%7D%20#0) ,a Regional Proposal Network [](https://www.codecogs.com/eqnedit.php?latex=%20%5Ctheta_%7Brpn%7D%20#0),two detection heads as [](https://www.codecogs.com/eqnedit.php?latex=%5Ctheta_%7Bd%7D%5E%7B1%7D%2C%20%5Ctheta_%7Bd%7D%5E%7B2%7D#0) 

## 第一步的启发：

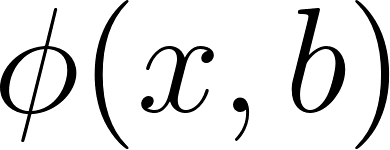
如果一个bbox是标注正确的，则两个不同的分类器会输出相同的预测结果，同时两个不同的分类器会对于背景类输出比较低的预测分数

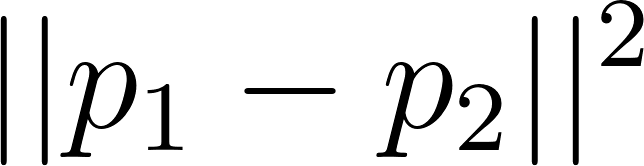
## First-step: CA-BBC(class-agnostic bounding box correction)

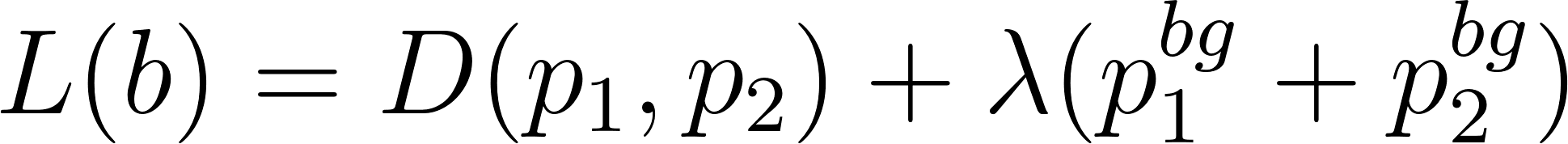


第一步专注于检测bbox的噪声

Aim: decouple bbox noise from label noise and optimize the noisy ground-truth bbox regardless of its class label

对于给定的图像x, backbone首先会提取x的卷积特征图。对于x中每一个方框b，使用RoI-pooling提取固定大小的特征[](https://www.codecogs.com/eqnedit.php?latex=%5Cphi(x%2Cb)#0),然后放入两个detection head中作softmax预测，从所有类别(包含背景图类)中得到两个预测结果向量p1, p2，使用欧式距离表示两个结果的差异性:

D(p1,p2) = [](https://www.codecogs.com/eqnedit.php?latex=%20%7C%7Cp_%7B1%7D%20-%20p_%7B2%7D%7C%7C%5E%7B2%7D#0) 把这个距离作为一个损失函数，若最小化会使得两个head预测一致类别，但是会趋向于使得box的位置移动到背景上，添加规范项使得背景类得到的分数最小:

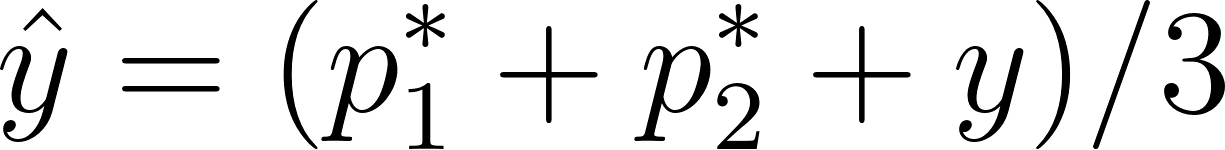
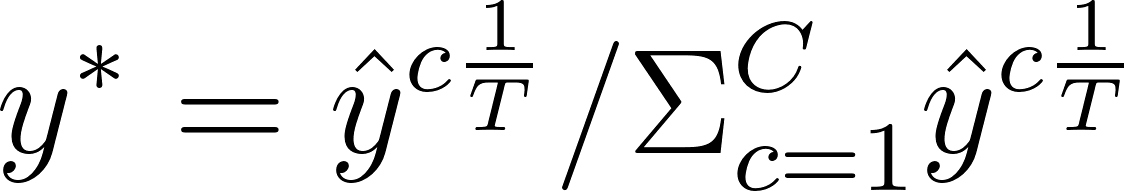
[](https://www.codecogs.com/eqnedit.php?latex=%20L(b)%20%3D%20D(p_%7B1%7D%2C%20p_%7B2%7D)%20%2B%5Clambda(p_%7B1%7D%5E%7Bbg%7D%20%2Bp_%7B2%7D%5E%7Bbg%7D)#0)

b根据L进行梯度下降就行(存在问题： 这里b属什么类型的数据变量？)

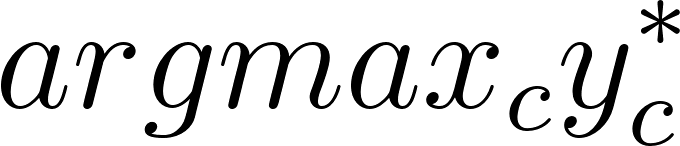
如果存在多标的假阳性问题(两个head预测的bg类分数均大于0.9)，通过reject模块判断并移除

## Second-step: class-specific bbox refinement

### Soft label correction

对于每一个提取的RoI特征并经过第一步的训练，两个分类头具有两个预测结果p1\*和p2\*，添加一个新变量y，用one-hot的形式代表ground truth,创造soft label [](https://www.codecogs.com/eqnedit.php?latex=%5Chat%7By%7D%20%3D%20(p_%7B1%7D%5E%7B*%7D%20%2B%20p_%7B2%7D%5E%7B*%7D%20%2By)%20%2F%203#0), 同时使用sharpening function优化softmax使得模型置信度预测[](https://www.codecogs.com/eqnedit.php?latex=y%5E%7B*%7D%20%3D%20%5Chat%7By%7D%5E%7Bc%5Cfrac%7B1%7D%7BT%7D%7D%20%2F%20%5CSigma_%7Bc%3D1%7D%5E%7BC%7D%5Chat%7By%7D%5E%7Bc%5Cfrac%7B1%7D%7BT%7D%7D#0), 其中c对应类别分数

### Class-specific bbox refinement

两个预测头预测了两组per-class boundng box regression offsets, t1和t2, 让c代表soft label中最高分数的类别c\* = [](https://www.codecogs.com/eqnedit.php?latex=argmax_%7Bc%7Dy_%7Bc%7D%5E%7B*%7D#0)。通过调整回归偏差，重新调整bbox，取偏差为两字预测偏差中关于c\*的均值

Training Details: 使用faster-RCNN的框架训练，并应用了co-teaching的训练方法

# OA-MIL(Object-Aware Multiple Instance approach)

Reference:<https://arxiv.org/pdf/2207.09697.pdf>

Open source: <https://github.com/cxliu0/OA-MIL>

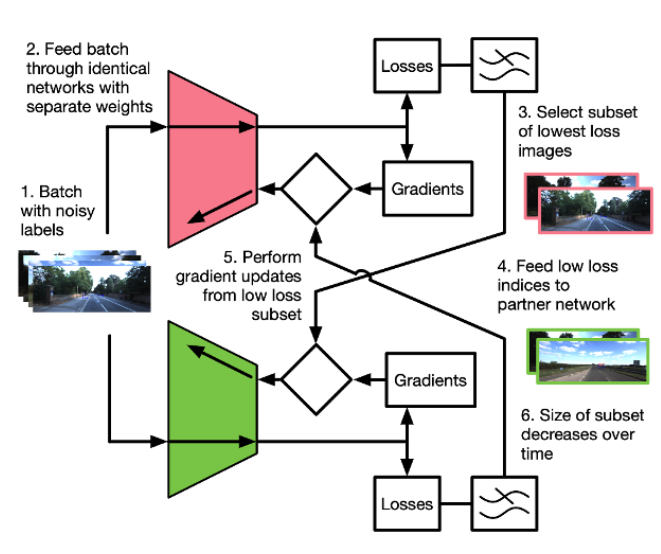
这个方法认为类别的预测作为ground-truth比标定的方框作为ground-truth更为准确，利用预测的类别去反向判定框的位置

**Focus on Noise-Robust Model Training**

# Co-teaching in Object Detection

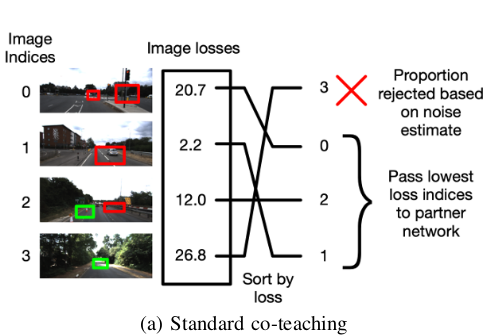
Reference: <https://arxiv.org/pdf/1905.07202.pdf>

source: <https://github.com/bhanML/Co-teaching> (image classification)



图像分类中的co-teaching方法：

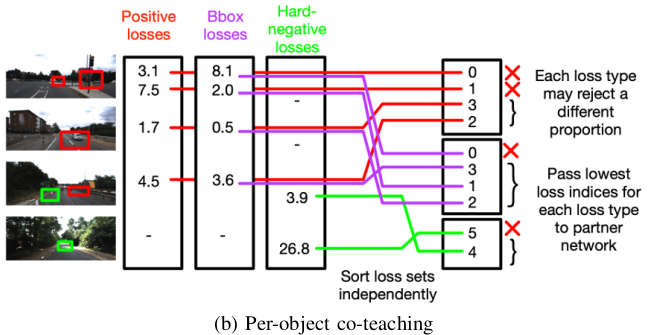
* 取用一个batch的数据
* 放到两个相同网络独立权重的模型中正向传播
* 从一个subset大小中选择最小的loss总和的images
* 把两个模型得到的loss相互交换进行训练，权重根据梯度下降依次更新
* Subset的大小随着时间减少



目标检测中的co-teaching方法：

梯度更新时，图像分类的方法是基于一整张图片的，而目标检测的方法是基于每一个目标的

对于一个object进行loss设计可以直接求和（如上图），文中由于采用SSD loss，所以将loss分级成positive lossses, bbox losses 和 hard-negative losses



对于每一个loss都有一部分比例存在于噪声中，分开取small 有利于每一种loss都被模型训练，从而能使得模型对各种噪声鲁棒而不是对于某一主要噪声鲁棒。