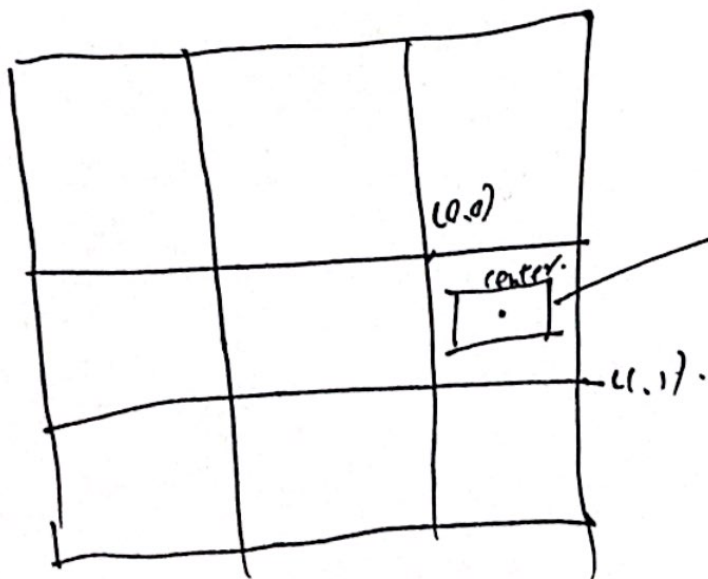


Topo. 算法

将图分解成网络 每个网长:

$\left\{ \begin{array}{l} pc \\ bx \\ by \\ bh \\ bw \\ c_1 \\ c_2 \\ c_3 \end{array} \right\}$

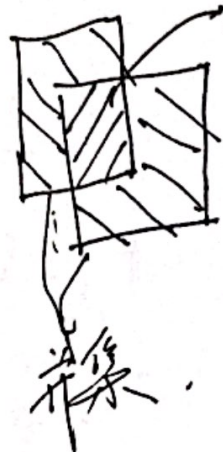
$\rightarrow$  是否有物体  
 $\rightarrow$  边界信息 (boundary box) 而不是网络  
 $\rightarrow$  物体种类



$y = \left\{ \begin{array}{l} 1 \\ 1.4 \\ 1.3 \\ 1.8 \\ 1.5 \\ 9 \\ 1 \\ 0 \end{array} \right\}$

$\rightarrow$  相对于网络来定位

交集



交集

$$IoU: \text{交集} = \frac{\text{交集}}{\text{并集}}$$

$$"correct" = IoU \geq 0.5$$

调整

非最大值抑制: 不足最大预测值的 boundary box 会被抑制

$$\text{if } p_c \leq 0.6$$

Prs card boundary boxes whose  $p_c \leq 0.6$

在剩下的边界框中取最大的  $p_c$  的 boundary box

抛弃低  $IoU$  的 ~~low~~ boundary box

## Anchor Boxes

处理各个目标出现在同一网络之中的问题, 所以根据 Anchor Boxes 对

物体进行区分

训练集中标:  $y = \begin{cases} p_{c1} \\ b_{x1} \\ b_{y1} \\ \vdots \end{cases}$  Anchor Box 1

$\begin{cases} p_{c2} \\ b_{x2} \\ b_{y2} \\ \vdots \end{cases}$  Anchor Box 2

$P_2(\text{object})$  { 有目标:  $Pr=1$   
无目标:  $Pr=0$

$$\therefore \text{Confidence} = P_2(\text{object}) * IOW$$

置信度

损失:

{ boundary box 损失  
Confidence 损失  
classes 损失.



# Yolo V~ (Yolo 9000)

1. BN层: (Batch Normalization)

在Conv层后加入BN层

2. Higher resolution classifier. 更大尺寸的特征分类器

3. Convolutional with Anchor boxes

4. Dimension clusters

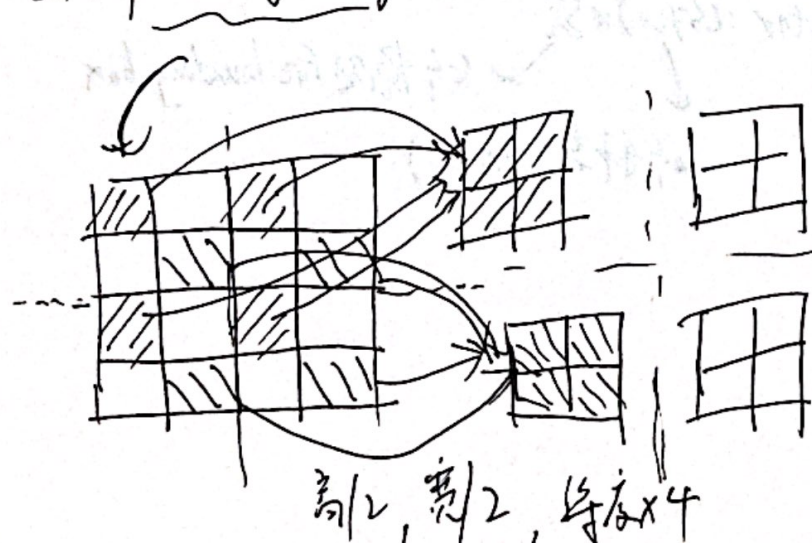
priors (相当于 Anchor box, default box) 使用 k-means 算法生成

5. Direct location prediction.

对 Anchor priors 的偏移量加以限制, 使其限制在 grid cell 内 提高边界框预测效率

6. Fine-Grained Features

通过 pass through layer 实现将层特征与高层特征融合。



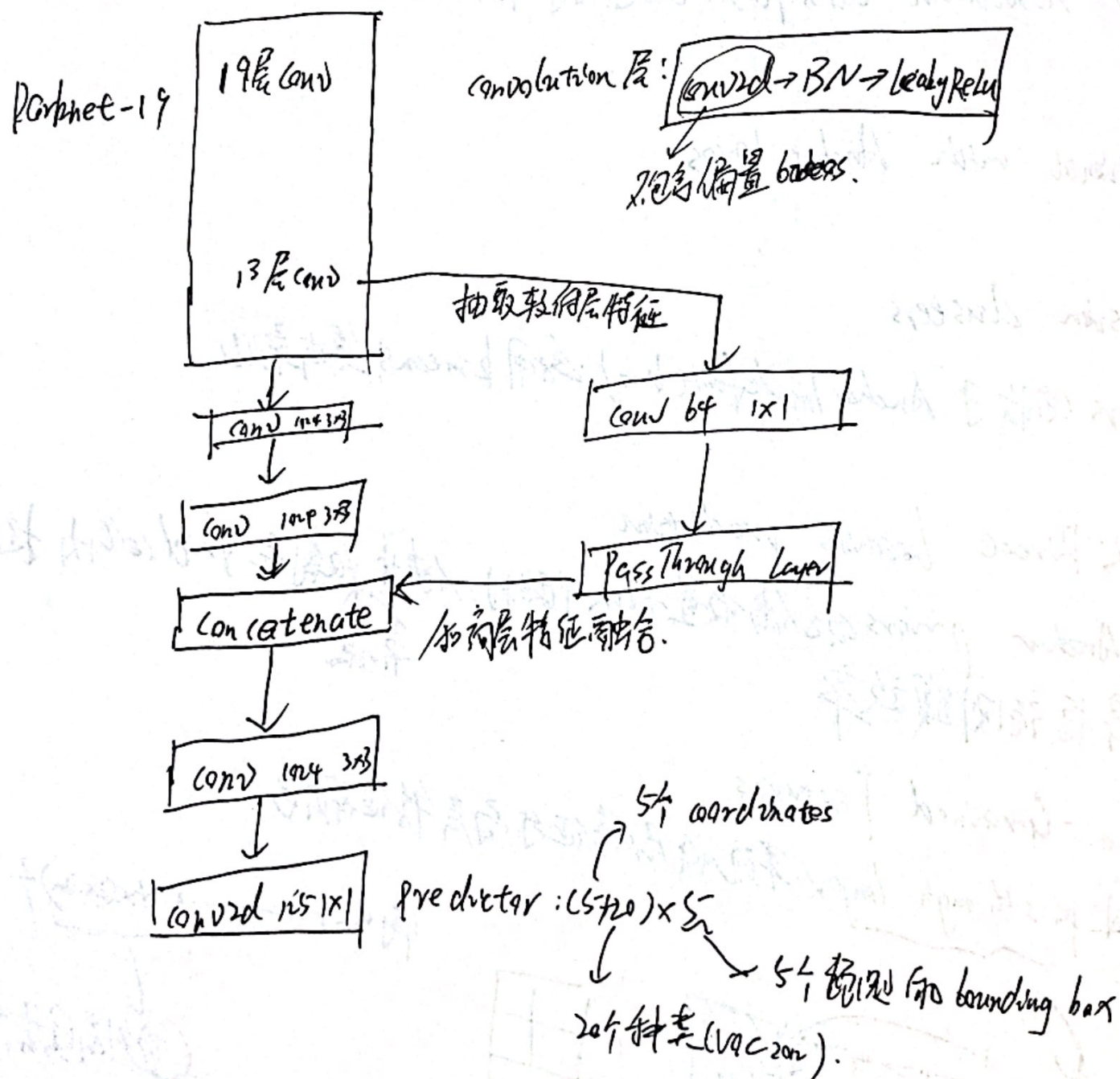
将  $4 \times 4 \rightarrow 4 \times (2 \times 2)$

加到高层特征  
深度方向拼接

Multi-scale training 多尺度训练.

每经过10个 batch 后 进行尺度变换输入.

Yolo v2 的 backbone: Darknet-19: 19层 Conv





Yolo v3:

将 backbone 由 darknet-19 改为 darknet-53.

darknet-53: 1. 不设置 pooling 层. 下采样使用 conv 实现.

2. 卷积核个数变少. 提升特征提取速度.