

# 基于深度学习的物体检测 - 作业 3

peng00bo00

August 23, 2020

1. 本次作业是对 FPN + Faster RCNN 代码进行调试, 现将代码核心内容总结如下:

结合 FPN 的 Faster RCNN 主要由 3 部分组成, 分别为 FPN、RPN 以及 RCNN

(a) FPN 负责生成特征金字塔, 计算流程为:

- i. 输入图像经过预处理后经过一个 backbone 网络 (代码中为 ResNet) 生成一个 4 层的特征金字塔
  - ii. 按照自上至下的顺序通过插值将上一层的特征图扩大 2 倍, 并与当前层的特征图进行融合
  - iii. 对最顶层的特征图进行一次 max pooling 加到特征金字塔中, 形成一个 5 层的特征金字塔
- FPN 最后得到一个包含 5 层的特征金字塔, 后面的 RPN 以及 RCNN 都会用到这个金字塔来进行计算。

(b) RPN 负责生成候选区域 (proposal):

- i. 根据输入的特征金字塔来生成对应特征图的锚框 (anchor)
- ii. 在不同尺度的特征图上利用卷积运算来计算每个锚框对应的置信度和偏移量
- iii. 将真实的标注 (Ground Truth) 映射到锚框上:
  - A. 计算每个锚框和每个标注的最大 IoU, 形成一个匹配矩阵 (match quality matrix)
  - B. 每个锚框与其具有最大 IoU 的标注进行匹配, 记录下对应的标注号和 IoU
  - C. 按照锚框具有的最大 IoU 将每个锚框划分为 3 类:  $\text{IoU} > 0.7$  的为前景 (正样例);  $\text{IoU} < 0.3$  的为背景 (负样例); 其余的忽略;
- iv. 按照锚框的类型进行采样, 得到 128 个包含前景的锚框和 128 个包含背景的锚框并将其余锚框标记为忽略
- v. 对采样得到的正负样例锚框按照置信度计算二分类损失并记录到 loss 中
- vi. 对采样得到的正样例锚框按照偏移量计算偏移量损失并记录到 loss 中
- vii. 将锚框的偏移量作用到锚框的位置上, 更新锚框的位置和大小
- viii. 基于锚框的置信度通过 NMS 筛选出置信度最大的前 1000 个锚框作为最终的候选区域

(c) RCNN 负责判断 ROI 的类别:

- i. 把真实的标注添加到 proposal 中
- ii. 计算每个 proposal 与真实标签的 IoU 并进行匹配, 按照 proposal 具有的最大 IoU 将每个锚框划分为 2 类:  $\text{IoU} \geq 0.5$  的为前景 (正样例);  $\text{IoU} < 0.0$  的为背景 (负样例);
- iii. 把实际标签的类别映射到 proposal 上
- iv. 对 proposal 进行采样, 得到 512 个包含各类标签的 proposal 其中前景最多占 25%
- v. 根据每个采样后 proposal 的大小计算所对应的尺度
- vi. 在每个 proposal 对应尺度的特征图上利用 ROI Align 将每个 proposal 规范化为指定的尺寸 ( $7 \times 7 \times 256$ )
- vii. 利用一个全连接网络计算规范化后 proposal 的特征以及 proposal 对应的类别和偏移量损失, 并将损失记录到 loss 中

最终对 loss 中各项进行求和并通过优化器来更新整个模型的参数