## 介绍

YOLO是一种目标检测算法，以其高速度和实时性著称。

其核心思想是将目标检测视为回归问题，通过单次前向传播即可完成对图像中所有目标的定位和分类。

## 原理

### 一、核心原理与步骤

1. 图像网格划分

YOLO将输入图像划分为S\*S的网格（例如YOLOv1采用7×7网格）。每个网格单元负责检测中心点落在该区域内的目标，并预测多个边界框（Bounding Box）。

2. 边界框与类别预测

每个网格单元预测B个边界框，每个框包含以下信息：（x,y,w,h,confidence）

- 坐标：中心点相对网格的偏移量(x, y)和归一化的宽高(w, h)（相对于整张图像的比例）。

- 置信度（Confidence）：表示框内存在目标的概率，并与预测框和真实框的交并比（IoU）相乘计算得出。

除此之外，还有如下信息：

**-** 类别概率：每个网格预测C个类别的概率，最终类别置信度为类别概率与边界框置信度的乘积

**-** 返回值：在预测完成后，YOLO会输出一个S\*S\*（B\*5+C）的张量

3. 置信度阈值与非极大值抑制（NMS）

- 阈值过滤：去除置信度低于设定阈值的边界框，减少冗余检测。

- 非极大值抑制：对同一目标的多个重叠框，仅保留置信度最高的框。通过计算IoU剔除重复检测结果。

### 网络结构与训练评价

1. 网络架构

- 骨干网络：早期YOLO采用GoogLeNet或DarkNet作为特征提取器，后续版本如YOLOv3使用DarkNet-53，通过卷积层CNN逐步提取特征并生成预测张量。

- 输出张量：最终输出为S\*S\*（B\*5+C）。

2. 训练评价：损失函数

YOLO的损失函数由三部分组成：

- 边界框坐标损失：计算预测边界框的中心坐标（x,y）和宽高（w,h）与真实值的误差

- 置信度损失：计算预测的置信度与真实值的误差。

- 类别损失：计算预测的类别概率与真实类别的误差。

## YOLO v1代码实现

<https://e.coding.net/xucancan1/yolov1/YOLOv1.git>

## 参考文档

<https://blog.csdn.net/hgnuxc_1993/article/details/116945869>

<https://www.ctyun.cn/developer/article/435054236770373>

<https://www.cnblogs.com/fariver/p/7446921.html>

<https://bbs.huaweicloud.com/blogs/297764>

<https://www.cnblogs.com/fariver/p/7446921.html>