# 阅读纲要

## 1 自己的总结、评价以及应用

本文提出了一种one step的object detection模型/算法——YOLO：You Only Look Once，相对于two step模型/算法，其整体性能是Fast但是accurancy较低。具体而言，其优点如下：

①YOLO is extremely fast.

②YOLO reasons globally about the image when making predictions.

③YOLO learns generalizable representations of objects.

该方法将object detection of an image视为a single regression problem，即We reframe object detection as a single regression problem, straight from image pixels to bounding box coordinates and class probabilities.即：

A single convolutional network simultaneously predicts multiple bounding boxes and class probabilities for those boxes.

## 2 文章的主要问题（abstract、疑问句中）

## 3 结论（abstract以及conclusion中）

## 4 思路脉络（小标题中的关键句）

1. Introduction
2. Unified Detection

重点就是将每张image划分为S\*S个grid cell，Each grid cell predicts B bounding boxes and confidence，并且Each bounding box consists of 5 predictions: x, y, w, h,and confidence.以及Each grid cell also predicts C conditional class probabilities, Pr(Class(i)|Object).

* 1. Network Design

Network的architecture：

CONV layer＋FC layer

We implement this model as a convolutional neural network and evaluate it on the PASCAL VOC detection dataset[9]. The initial convolutional layers of thenetwork extract features from the image while the fully connected layers predict the output probabilities and coordinates.

Our network has 24

convolutional layers followed by 2 fully connected layers.Instead of the inception modules used by GoogLeNet, we simply use 1 × 1 reduction layers followed by 3 × 3 convolutional layers, similar to Lin et al [22].

* 1. Training
  2. Inference
  3. Limitations of YOLO

YOLO存在的disadvantages：

①每个grid cell的bounding box数量有限，且很难（struggle to）识别小物体

YOLO imposes strong spatial constraints on bounding box predictions since each grid cell only predicts two boxes and can only have one class. This spatial constraint limits the number of nearby objects that our model can predict. Our model struggles with small objects that appear in groups, such as flocks of birds

②很难将其推广到具有新的或不寻常的宽高比或配置的对象（灵活性不高）

③our loss function treats errors the same in small bounding boxes versus large bounding boxes.

1. Comparison to Other Detection Systems
2. Experiments
3. Real-Time Detection In The Wild
4. Conclusion

We introduce YOLO, a unified model for object detection. Our model is simple to construct and can be trained directly on full images.

## ５ 难理解点