

MMDetection论文阅读笔记

MMDetection是商汤联合港中文开发的一个基于pytorch的深度学习目标检测框架，框架中提供了目前主流的一些目标检测的模型。最近准备学习一下MMDetection这个目标检测框架，这篇文章是MMDetection论文的阅读笔记。

论文地址：[MMDetection: Open MMLab Detection Toolbox and Benchmark](https://arxiv.org/abs/1802.04268)

MMDetection论文阅读笔记

1. 论文摘要
2. MMDetection框架整体介绍
 - 2.1 MMDetection框架支持的模型
 - 2.2 模型架构
 - 2.3 训练pipeline
 - 2.4 其他的模块

1. 论文摘要

我们提出了一个目标检测工具箱MMDetection，这个工具箱包含了大量的目标检测与实例分割的方法以及一些相关的模块与组件。这个工具箱开始于2018年COCO检测挑战赛冠军队伍MMDet的代码库，逐渐进化成一个包含流行的目标检测算法以及现在常用模块的统一的目標检测平台。MMDetection不仅包含了训练与推理的代码，而且提供了200多种预训练模型（现在好像已经有300多种了）。我们相信这个框架将会慢慢变成一个最完善的目标检测工具箱。在这片论文中，我们介绍了这个工具箱很多的特性，除此之外我们也对不同的方法，不同的组件和他们的超参数进行了基准研究。我们希望这个提供弹性工具的工具箱能够服务于逐渐发展庞大的研究社区，让他们可以方便的重新实施已经存在的算法，并且研究开发自己新的检测器。代码地址：<https://github.com/open-mmlab/mmdetection>。

2. MMDetection框架整体介绍

1. 模块化设计

将目标检测算法解耦为不同的组件，研究者可以通过组合不同的模块快速开发出自己的目标检测框架

2. 支持很多目标检测框架

3. 高效

所有的bbox与mask操作均在gpu上进行处理，速度更快（相比较Detectron, maskrcnn-benchmark,simpleDet这几个框架）。

4. state of art

MMDet的队伍获得了COCO检测挑战赛2018年的冠军。这个工具箱就是基于MMDet的代码库开发的。

2.1 MMDetection框架支持的模型

Table 1: Supported features of different codebases. “✓” means officially supported, “*” means supported in a forked repository and blank means not supported.

	MMDetection	maskrcnn-benchmark	Detectron	SimpleDet
Fast R-CNN	✓	✓	✓	✓
Faster R-CNN	✓	✓	✓	✓
Mask R-CNN	✓	✓	✓	✓
RetinaNet	✓	✓	✓	✓
DCN	✓	✓	✓	✓
DCNv2	✓	✓		
Mixed Precision Training	✓	✓		✓
Cascade R-CNN	✓		*	✓
Weight Standardization	✓	*		
Mask Scoring R-CNN	✓	*		
FCOS	✓	*		
SSD	✓			
R-FCN	✓			
M2Det	✓			
GHM	✓			
ScratchDet	✓			
Double-Head R-CNN	✓			
Grid R-CNN	✓			
FSAF	✓			
Hybrid Task Cascade	✓			
Guided Anchoring	✓			
Libra R-CNN	✓			
Generalized Attention	✓			
GCNet	✓			
HRNet	✓			
TridentNet [17]				✓

2.2 模型架构

虽然不同模型的架构不同，但是其基本的结构可以分为以下几个部分：

Backbone：将图片转换为feature map，例如不包含全连接层的resnet50

Neck：Neck是连接backbone与head的一部分网络，主要工作就是对backbone得到的feature map进行重构融合，代表为FPN

DenseHead：在feature map上，执行一个密集的定位操作，包含了anchorHead与anchorFreeHead。

ROIExtractor：利用类似ROIpooling的操作从单一或者多层特征图中提取感兴趣区域的操作

ROIHead：利用ROI提取器得到的ROI作为输入，然后进行bbox分类回归或者mask预测。

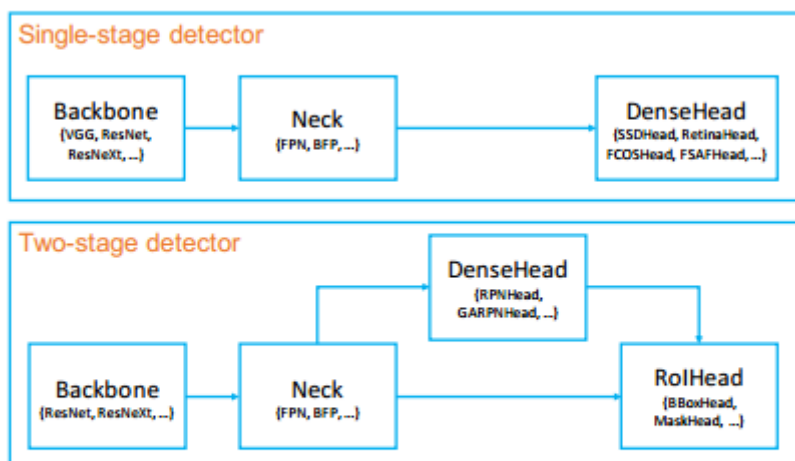


Figure 1: Framework of single-stage and two-stage detectors, illustrated with abstractions in MMDetection.

2.3 训练pipeline

采用hook机制设计了一个统一的训练pipeline，这种设计不仅在目标检测中可以使用，在其他的计算机视觉任务中也可以使用。

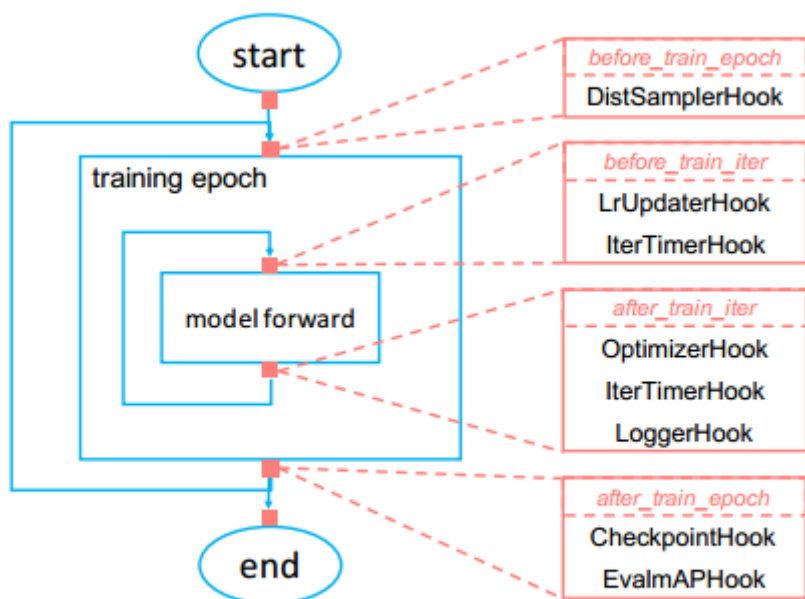


Figure 2: Training pipeline.

2.4 其他的模块

框架还提供了多尺度训练，多尺度预测的方案

提供了多种loss

多种neck方案

多种nms的策略

总之这个框架几乎涵盖了所有目标检测中常用的一些策略及模型。