

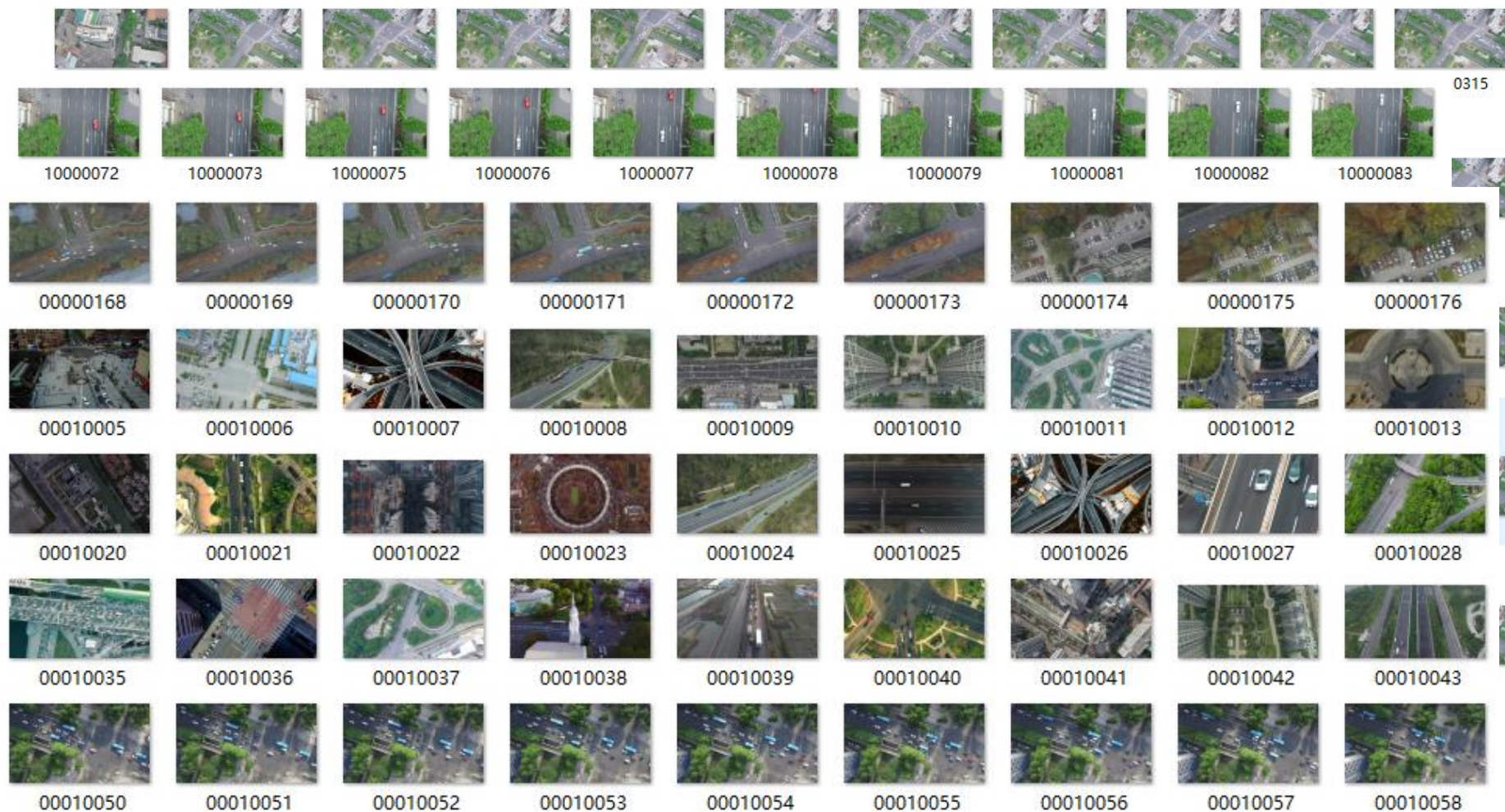


航空杯竞赛成果展示

HKCSEU124



数据集概览





Retinanet

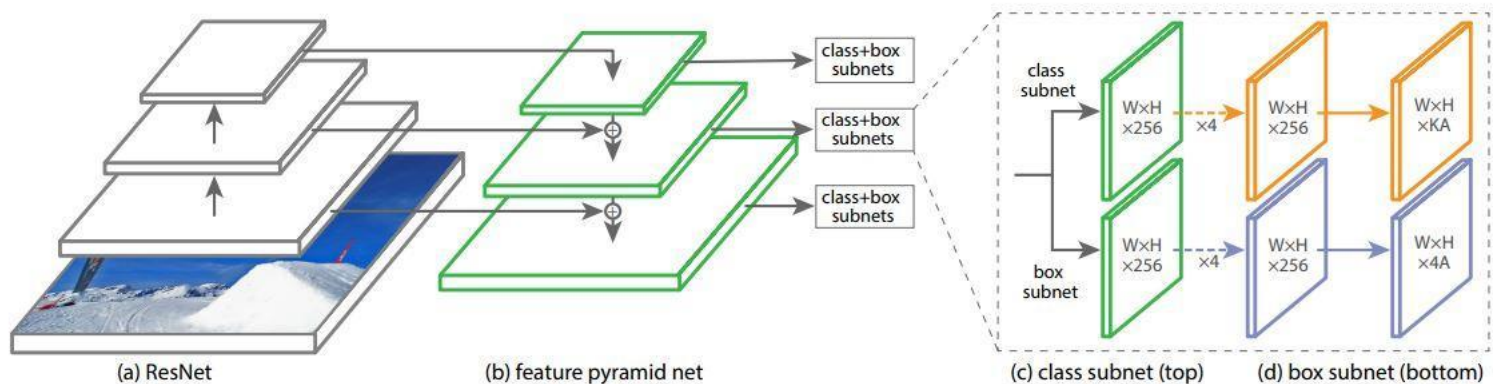


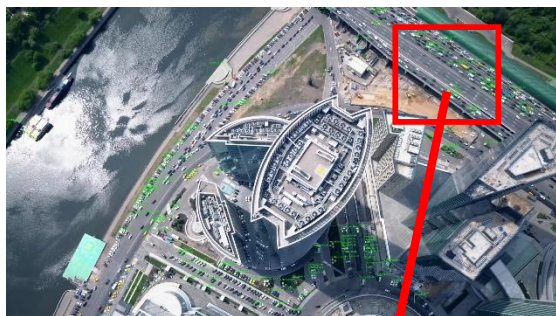
Figure 3. The one-stage **RetinaNet** network architecture uses a Feature Pyramid Network (FPN) [20] backbone on top of a feedforward ResNet architecture [16] (a) to generate a rich, multi-scale convolutional feature pyramid (b). To this backbone RetinaNet attaches two subnetworks, one for classifying anchor boxes (c) and one for regressing from anchor boxes to ground-truth object boxes (d). The network design is intentionally simple, which enables this work to focus on a novel focal loss function that eliminates the accuracy gap between our one-stage detector and state-of-the-art two-stage detectors like Faster R-CNN with FPN [20] while running at faster speeds.

RetinaNet是一种one-stage目标检测方法。本质上是Resnet + FPN + 两个FCN子网络。另外通过Focal Loss抑制那些容易分类样本的权重，将注意力集中在那些难以区分的样本上，有效控制正负样本比例，防止失衡。

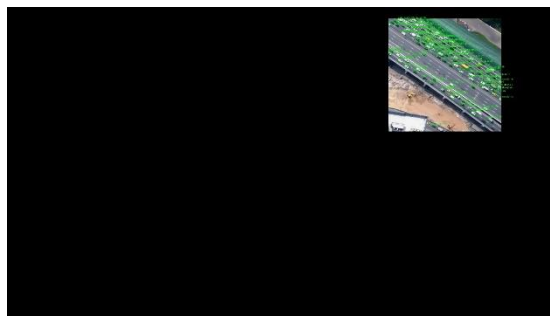


创新点1-窗口分割

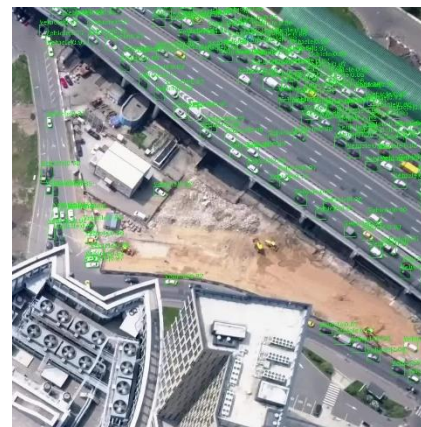
□ 三种输入效果对比



原图



窗口之外部分填充0得到的图



窗口截取出的图



\wedge

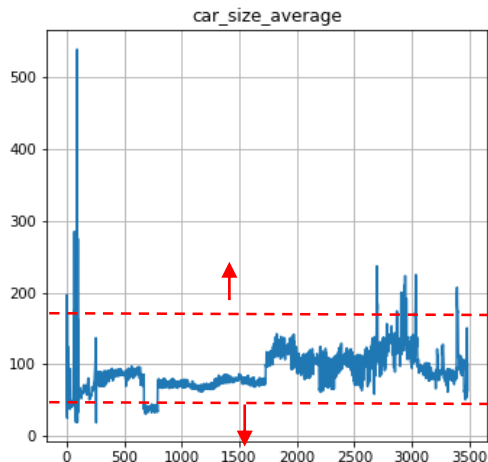
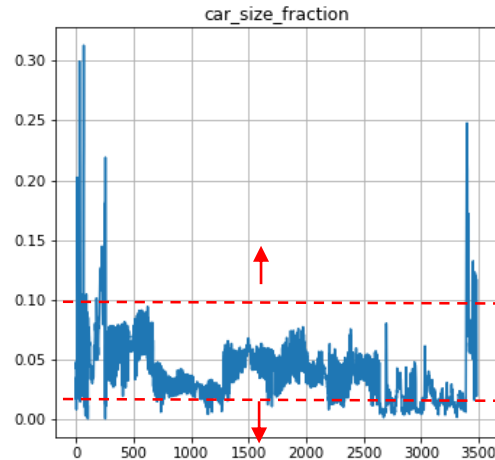
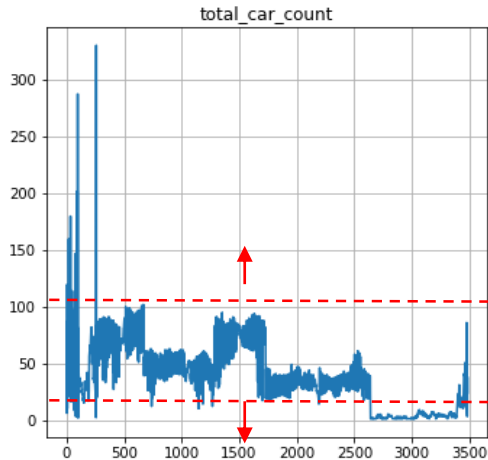


\approx





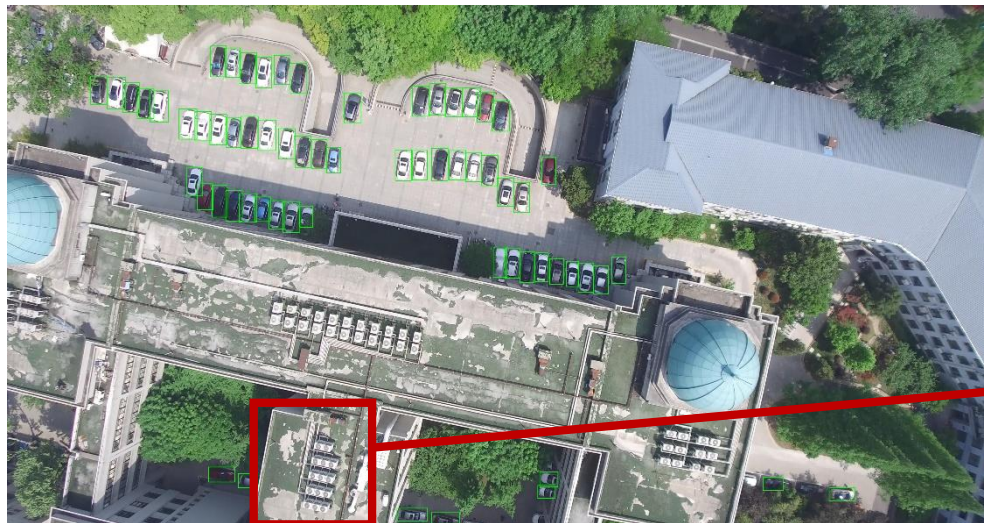
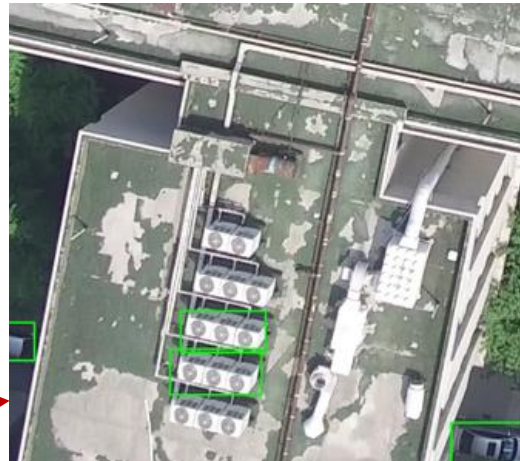
创新点2-难例重采样



- 每张图车的数量
- 每张图有车区域占比
- 每张图单辆车平均大小

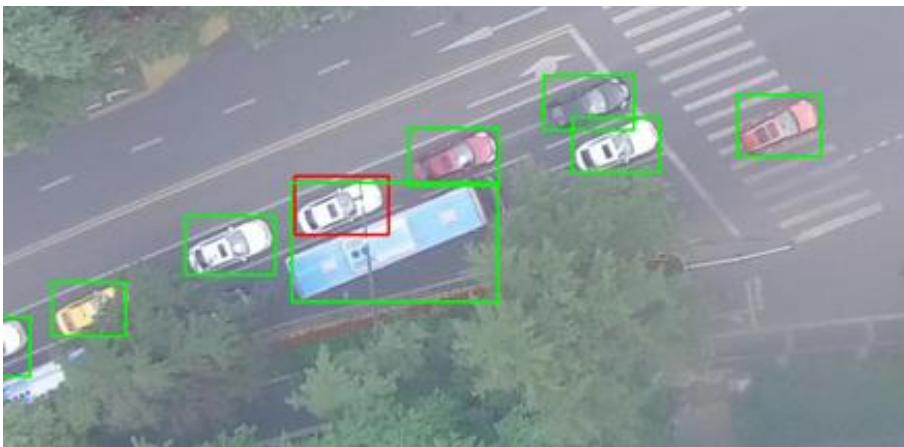
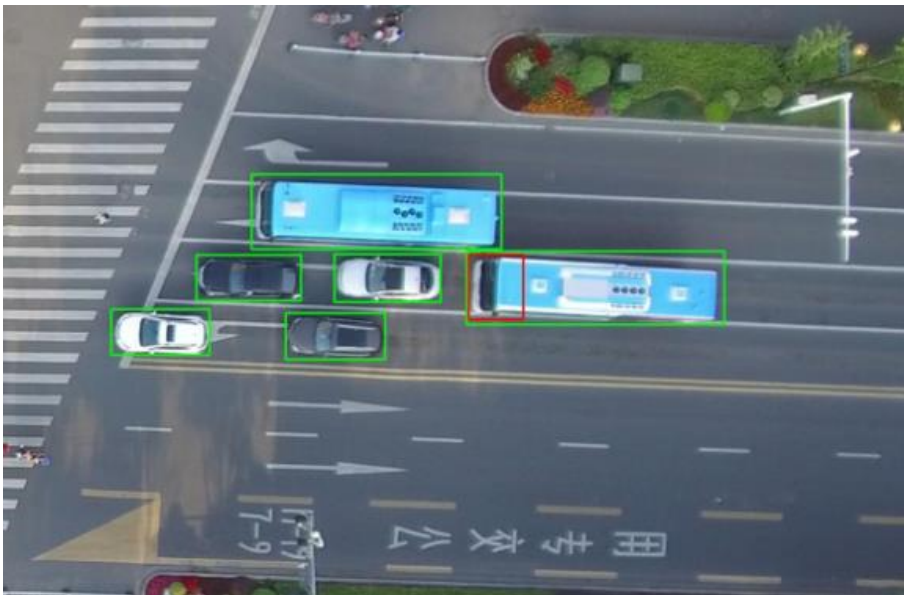


创新点2-难例重采样



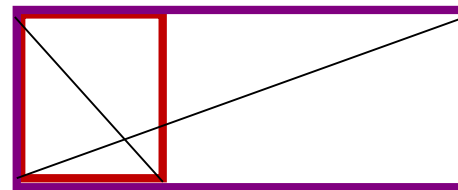


创新点3-结果的后处理

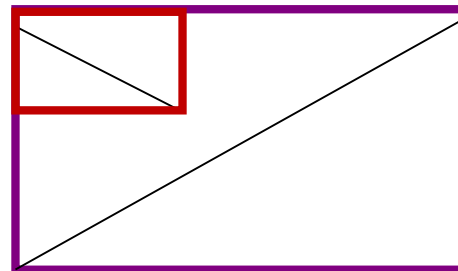


规则：

- 左上角的点比较近
- 小框的三边与大框重合



False



True

前提：

- 数据集是俯拍图



其他技巧

FPN（特征金字塔网络）	基本不增加原有模型计算量情况下，大幅度提升了小物体检测的性能
Backbone	选用了Resnext101,其在ImageNet上获得了85.4% top-1 准确率的参数
Focal Loss	通过抑制那些容易分类样本的权重，将注意力集中在那些难以区分的样本上，有效控制正负样本比例，防止失衡
数据增强	增强鲁棒性，减少模型的过拟合
多尺度训练	提升了模型对于尺度差异大的场景的适应性
Guided anchor	通过图像特征来指导 anchor 的生成，提高proposal 选取的质量



结果

- 随机拆分全部数据为训练集和验证集
- train set: 3484; val set: 871
- 最终所得模型处理1080p图像的平均时间为93.7ms（GPU）
- 训练与测试均在mmdetection框架上完成

Resnet50	Resnext101	Ratio unfix	Guided anchoring	图像左右切割	图像固定窗口切割	难例重采样训练	Result F1-score
√							90.9
√		√					91.7
	√	√					93.4
	√	√	√				94.2
	√		√	√			95.9
	√		√		√		97.5
	√		√		√	√	98.3



Q&A
